
 WISE (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Dec 18 18:14:18 1998: MasPar time 50.54 Seconds
 Tabular output not generated. 935.076 Million cell updates/sec

Title: >US-08-951-733-14
 Description: (1-949) from US08951733.pep (1 of 2)
 Perfect Score: 7113
 Sequence: 1 HASGORCVLRITWEALAPAT.....PYDEALGCTAFVQMPAHGL 949

Scoring table: PAM 150
 Gap 11

Searched: 165420 segs, 49795644 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: sptrembl6
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
 9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 53.273; Variance 101.693; scale 0.524

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	6954	97.8	1132	4	014746	TELOMERASE REVERSE TRA	0.00e+00
2	6948	97.7	1132	4	014783	TELOMERASE CATALYTIC S	0.00e+00
3	4087	57.5	1122	11	070372	TELOMERASE REVERSE TRA	0.00e+00
4	675	9.5	989	3	013338	TELOMERASE REVERSE TRA	2.28e-108
5	672	9.4	988	3	013339	TELOMERASE REVERSE TRA	9.66e-108
6	432	6.1	1031	5	000939	TELOMERASE SUBUNIT P12	2.78e-58
7	424	6.0	67	11	035432	TELOMERASE CATALYTIC S	1.11e-56
8	314	4.4	884	3	006163	CHROMOSOME XII COSMID	3.84e-35
9	147	2.1	296	14	069118	HYPOTHETICAL PROTEIN (6.53e-06
10	142	2.0	234	4	000600	PAROTID 'O' PROTEIN (F	3.60e-05
11	141	2.0	276	2	046612	IS 1222 GENE ONF-A AND	5.04e-05
12	140	2.0	373	2	054226	POLYKETIDE SYNTHASE (F	7.05e-05
13	143	2.0	574	3	036027	WISKOTT-ALDRICH SYNDRO	2.56e-05
14	136	1.9	316	4	098076	HOMEOBOX PROTEIN (FRAG	2.67e-04
15	132	1.9	317	11	062103	PROLINE RICH PROTEIN P	9.90e-04
16	133	1.9	539	5	002123	SIMILARITY TO COLLAGEN	7.15e-04
17	135	1.9	585	14	041935	HYPOTHETICAL 60.2 KD P	3.71e-04
18	136	1.9	897	11	070495	PLENTY-OF-PROLINES-101	9.90e-04
19	132	1.9	924	14	099307	LE3 PROTEIN.	9.90e-04
20	125	1.8	227	11	062107	PROLINE-RICH SALIVARY	9.25e-03

21	129	1.8	264	14	085301	HOMOLOGUE OF RETROVIRA	2.60e-03
22	131	1.8	300	11	061888	PROLINE RICH PROTEIN.	1.37e-03
23	129	1.8	309	4	004118	SALIVARY PROLINE-RICH	2.60e-03
24	125	1.8	503	4	043516	WASP INTERACTING PROTE	9.25e-03
25	128	1.8	527	2	035913	BETA-KETOACYL SYNTHASE	3.58e-03
26	130	1.8	1236	2	006264	HYPOTHETICAL 131.9 KD	1.89e-03
27	126	1.8	3247	14	065553	UL36	6.75e-02
28	118	1.7	203	5	P1497	CODED FOR BY C. ELEGAN	7.98e-02
29	118	1.7	236	14	085020	SUBSTRATE OF THE PROTE	7.98e-02
30	118	1.7	322	2	053180	MP13030 NODDI GENE IN	7.98e-02
31	118	1.7	405	2	068033	EXONUCLEASE SBCD HOMOL	7.98e-02
32	118	1.7	437	2	052256	FLHE	7.98e-02
33	123	1.7	439	10	042421	CHITINASE PRECURSOR.	1.73e-02
34	124	1.7	464	10	041645	EXTENSIN (FRAGMENT).	1.26e-02
35	123	1.7	494	4	015220	PRP-2 PROTEIN.	1.73e-02
36	118	1.7	526	14	085027	VIRAL PROTEINASE.	7.98e-02
37	122	1.7	657	10	022835	SF16 ISOLOG.	2.35e-02
38	124	1.7	820	4	060585	SER/ARG-RELATED NUCLEA	1.26e-02
39	115	1.6	106	14	041981	HYPOTHETICAL 9.9 KD PR	1.96e-01
40	117	1.6	197	5	017626	C0462.8.	1.08e-01
41	115	1.6	260	11	064306	PROLINE-RICH PROTEIN.	1.96e-01
42	116	1.6	295	11	007611	PROLINE-RICH PROTEOGLY	1.45e-01
43	115	1.6	297	4	016038	PRBIM PROTEIN (FRAGMEN	1.96e-01
44	115	1.6	382	4	000599	CONT.	1.96e-01
45	116	1.6	488	10	004528	SEQUENCE OF BAC F20P5	1.45e-01

ALIGNMENTS

RESULT	ID	1	PRELIMINARY;	PRT;	1132 AA.
AC	014746	01-JAN-1998 (TREMBLEL. 05, CREATED)			
DT	01-JAN-1998 (TREMBLEL. 05, LAST SEQUENCE UPDATE)				
DT	01-JUN-1998 (TREMBLEL. 06, LAST ANNOTATION UPDATE)				
DE	TELOMERASE REVERSE TRANSCRIPTASE.				
GN	HTRT.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
CC	EUTHERIA; PRIMATES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-KIDNEY;				
FX	MEDLINE; 97400623.				
RA	NANAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,				
RL	LINGNER J., HARLEY C.B., CECCH T.R.;				
DR	SCIENCE 277:955-959(1997).				
KW	EMBL; AF015950; G2330017; -				
SO	RNA-DIRECTED DNA POLYMERASE.				
SO	SEQUENCE 1132 AA; 126995 MW; 2DFBEDF3 CRC32;				
Query Match	97.8%;	Score 6954;	DB 4;	Length 1132;	
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;			
Matches	927;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Db	1	MPRAPRCRAVSLRSRREVLPLATFPRRLGPOGRRLVQSGDPAAPALVAOCLVCPW	60		
Qy	23	MPRAPRCRAVSLRSRREVLPLATFPRRLGPOGRRLVQSGDPAAPALVAOCLVCPW	82		
Db	61	DARPPAPSPFROVSCLELVAFLVQRLCEGAKNVLAFLGALDAGARGPEAFSTSVR	120		
Qy	83	DARPPAPSPFROVSCLELVAFLVQRLCEGAKNVLAFLGALDAGARGPEAFSTSVR	142		
Db	121	SYLPRTVTDALRSGAWGLLRVGGDDVLHILARCALFVLVAPSCAYQVCGPPLYQIGA	180		
Qy	143	SYLPRTVTDALRSGAWGLLRVGGDDVLHILARCALFVLVAPSCAYQVCGPPLYQIGA	202		
Db	181	ATQARPPPHASGPRRLRCCERAMNHVSVEAGVPLGLPAPGARRGSGSRSRLPLPKRRR	240		
Qy	203	ATQARPPPHASGPRRLRCCERAMNHVSVEAGVPLGLPAPGARRGSGSRSRLPLPKRRR	262		
Db	241	GAAPERTPVGSGMAHPGRTGSPDGRGCVVSPAPPAEATLEGALSTRSHPSVG	300		

```
QY 263 GAAPERPTVGGSWAHNGRTGSDRGFCVVSPPARPAEATSLGALSTGRSHSPVG 322
Db 301 ROHHAGPSTSRPPRPMDTCCPPVYATKHFLLSSGDKEQLRPSFLSSRLPSTLGARL 360
QY 323 ROHHAGPSTSRPPRPMDTCCPPVYATKHFLLSSGDKEQLRPSFLSSRLPSTLGARL 382
Db 361 VERIFLGSRRPMGTPRRRLPRLPQRYQMORPLFLELLGNHACPYGVLLTHCPLRAVT 420
QY 383 VERIFLGSRRPMGTPRRRLPRLPQRYQMORPLFLELLGNHACPYGVLLTHCPLRAVT 442
Db 421 PAAGVCAREKPOGSVAAPBEEDTDPRLVOLLROHSSPMOYGFVACLRRLVPPGLMS 480
QY 443 PAAGVCAREKPOGSVAAPBEEDTDPRLVOLLROHSSPMOYGFVACLRRLVPPGLMS 502
Db 481 RHNERRLRNTKFKFISLGNHAKLSLOELTWKMSVRDCAMLRSPGVCVPAAEHRLREI 540
QY 503 RHNERRLRNTKFKFISLGNHAKLSLOELTWKMSVRDCAMLRSPGVCVPAAEHRLREI 562
Db 541 LAKEFLHMLMSVYVELLSRFFVYETTFOKNRLFEYRKSVMSKLSIGIRHOLKRVOLRE 600
QY 563 LAKEFLHMLMSVYVELLSRFFVYETTFOKNRLFEYRKSVMSKLSIGIRHOLKRVOLRE 622
Db 601 LSEAEVQRHREARPALTLRLRFPKPDGLRPIVNDYVVGARTFRERKRAERLTSVKA 660
QY 623 LSEAEVQRHREARPALTLRLRFPKPDGLRPIVNDYVVGARTFRERKRAERLTSVKA 682
Db 661 LFSVLNTERARRRGLLGASVGLDDIHRAMRTFVLRYRADDPPELTFVYVDTGAYDTI 720
QY 683 LFSVLNTERARRRGLLGASVGLDDIHRAMRTFVLRYRADDPPELTFVYVDTGAYDTI 742
Db 721 PODRLTEVIASIIKPPONTYCVRRYAVVOKAAGHVRAKAFSHVSTLTDLPYMRQFAHL 780
QY 743 PODRLTEVIASIIKPPONTYCVRRYAVVOKAAGHVRAKAFSHVSTLTDLPYMRQFAHL 802
Db 781 QETSPLRDAAVYIEOSSSLNEASSGLFDVFLRFMCHAAVRIRGKSYVOCQIGIPOSILSTL 840
QY 803 QETSPLRDAAVYIEOSSSLNEASSGLFDVFLRFMCHAAVRIRGKSYVOCQIGIPOSILSTL 862
Db 841 LCSLCYDMDENKLFAGIRRGGLLLRLVDDFLVTPHLTHAKTFLRTLVYRGVPEYGCYVNL 900
QY 863 LCSLCYDMDENKLFAGIRRGGLLLRLVDDFLVTPHLTHAKTFLRTLVYRGVPEYGCYVNL 922
Db 901 RKTIVNFPVEDEALGTAFAVOMPAHGL 927
QY 923 RKTIVNFPVEDEALGTAFAVOMPAHGL 949

RESULT 2 PRELIMINARY; PRT; 1132 AA.
ID 014783;
AC 014783;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TELOMERASE CATALYTIC SUBUNIT.
GN HEST2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE: 97433088.
RA MEYERSON M., COUNTER C.M., EATON E.N., ELLISEN L.W., STEINER P.,
RA CADDLE S.D., ZIAUGRA L., BEIJERSBERGEN R.L., DAVIDOFF M.J., LIO Q.,
RA BACCHETTI S., HABER D.A., WEINBERG R.A.;
RL CELL 90:785-795(1997).
DR EMBL: AF018167: G2347129;
SQ SEQUENCE 1132 AA: 126937 MM: C1E5E2AF CRC32;

Query Match 97.7%; Score 6948; DB 4; Length 1132;
Best local similarity 99.9%; Pred. No. 0.00e+00;
Matches 926; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 23 MPAPRCRAVRSLLSRSHYREVLPATFVRLRIGPQGMRLVQRGDPAARFALVACLVCPW 82
Db 61 DARRPAPASFRVOSCKELVAVIOLRCEGAKNYLAFGALLDAGRGPPAFTSVR 120
QY 83 DARRPAPASFRVOSCKELVAVIOLRCEGAKNYLAFGALLDAGRGPPAFTSVR 142
Db 121 STLPNTVTALRGSGAMGLLRVYGDVLYHLLARCALVIVASGAYOVCPRLYOLA 180
QY 143 STLPNTVTALRGSGAMGLLRVYGDVLYHLLARCALVIVASGAYOVCPRLYOLA 202
Db 181 ATQARRPPHASGPRRLRGCERAMNHSVREAGVPLGIPAPGARRGGSASRSLLPRKPRR 240
QY 203 ATQARRPPHASGPRRLRGCERAMNHSVREAGVPLGIPAPGARRGGSASRSLLPRKPRR 262
Db 241 GAAPERPTVGGSWAHNGRTGSDRGFCVVSPPARPAEATSLGALSTGRSHSPVG 300
QY 263 GAAPERPTVGGSWAHNGRTGSDRGFCVVSPPARPAEATSLGALSTGRSHSPVG 322
Db 301 ROHHAGPSTSRPPRPMDTCCPPVYATKHFLLSSGDKEQLRPSFLSSRLPSTLGARL 360
QY 323 ROHHAGPSTSRPPRPMDTCCPPVYATKHFLLSSGDKEQLRPSFLSSRLPSTLGARL 382
Db 361 VERIFLGSRRPMGTPRRRLPRLPQRYQMORPLFLELLGNHACPYGVLLTHCPLRAVT 420
QY 383 VERIFLGSRRPMGTPRRRLPRLPQRYQMORPLFLELLGNHACPYGVLLTHCPLRAVT 442
Db 421 PAAGVCAREKPOGSVAAPBEEDTDPRLVOLLROHSSPMOYGFVACLRRLVPPGLMS 480
QY 443 PAAGVCAREKPOGSVAAPBEEDTDPRLVOLLROHSSPMOYGFVACLRRLVPPGLMS 502
Db 481 RHNERRLRNTKFKFISLGNHAKLSLOELTWKMSVRDCAMLRSPGVCVPAAEHRLREI 540
QY 503 RHNERRLRNTKFKFISLGNHAKLSLOELTWKMSVRDCAMLRSPGVCVPAAEHRLREI 562
Db 541 LAKEFLHMLMSVYVELLSRFFVYETTFOKNRLFEYRKSVMSKLSIGIRHOLKRVOLRE 600
QY 563 LAKEFLHMLMSVYVELLSRFFVYETTFOKNRLFEYRKSVMSKLSIGIRHOLKRVOLRE 622
Db 601 LSEAEVQRHREARPALTLRLRFPKPDGLRPIVNDYVVGARTFRERKRAERLTSVKA 660
QY 623 LSEAEVQRHREARPALTLRLRFPKPDGLRPIVNDYVVGARTFRERKRAERLTSVKA 682
Db 661 LFSVLNTERARRRGLLGASVGLDDIHRAMRTFVLRYRADDPPELTFVYVDTGAYDTI 720
QY 683 LFSVLNTERARRRGLLGASVGLDDIHRAMRTFVLRYRADDPPELTFVYVDTGAYDTI 742
Db 721 PODRLTEVIASIIKPPONTYCVRRYAVVOKAAGHVRAKAFSHVSTLTDLPYMRQFAHL 780
QY 743 PODRLTEVIASIIKPPONTYCVRRYAVVOKAAGHVRAKAFSHVSTLTDLPYMRQFAHL 802
Db 781 QETSPLRDAAVYIEOSSSLNEASSGLFDVFLRFMCHAAVRIRGKSYVOCQIGIPOSILSTL 840
QY 803 QETSPLRDAAVYIEOSSSLNEASSGLFDVFLRFMCHAAVRIRGKSYVOCQIGIPOSILSTL 862
Db 841 LCSLCYDMDENKLFAGIRRDGILLRLVDDFLVTPHLTHAKTFLRTLVYRGVPEYGCYVNL 900
QY 863 LCSLCYDMDENKLFAGIRRDGILLRLVDDFLVTPHLTHAKTFLRTLVYRGVPEYGCYVNL 922
Db 901 RKTIVNFPVEDEALGTAFAVOMPAHGL 927
QY 923 RKTIVNFPVEDEALGTAFAVOMPAHGL 949

RESULT 3 PRELIMINARY; PRT; 1122 AA.
ID 070372;
AC 070372;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
```


[illegible]

SEQ	SEQUENCE	1031 AA:	122562 MW:	21A885CD	CRC32:
Query Match	6.1%:	Score 432:	DB 5:	Length 1031:	
Best Local Similarity	25.2%:	Fred. No. 2.78e-36:			
Matches 119:	Conservative 122:	Mismatches 194:	Indels 38:	Gaps 27:	
Db	361	LINEFEYINLTKDFLTGR-NRKNQKVKYKVEVLNKH- LTHKKMLLEKINTREISMNOV	418		
QY	486	FYRACLRLVPGGLMGSHNRRRLRNTKRTISGKNAKLSLOELTW-KMSVRCANLRR	544		
Db	419	ETSAKHFFYPDHE-NIYVLKRLRMIEDVLSIRCFEYVTEOQSYSKTYRRKNIWD	477		
QY	545	SPGVGCAVAABHRLREELIAKFLHMSVYVELLRSEFYVTEETTFQKNRLFEYRKSYWS	604		
Db	478	VIMKMSI-ADLKEETLAEVOKEV-EEMKKSGLFAPGKRLIPKTTFRPI--MTF--NK	531		
QY	605	KLOSIGIQHKLRYOLRELSEAEVQHREARPAALTSRLREIPKPDGLRPIVNDYVGA	664		
Db	532	KIVNSDKRTKLTINTKLLNSHMLKTLKNMFDPDGFAPAFVNDVWKVKEEFVCK-WK	590		
QY	665	KTRERKRAERLTSYVKLFS--VLNERAR--PGLGASVGLDDIHRAMRTVLNRA	721		
Db	591	QVGQKLEFPATMDEICEQDSVNRKEISFEKTKLLSDSEVIMTAQLKRNIVIDSKN	650		
QY	722	QDPPELTFYVADVTGADTIPDRLEFVI-AS-IIRQ--NRYCV--RRYAV-Q-KA	772		
Db	651	FRKEMKDYFROKQFALLEGQYPTLESVLNENDNDINAKTLIVEAKORNYKKDNL	710		
QY	773	AHGVRKA-FKSHVSTL-DLQPMRF-V-AHQETSPLDADVAIVIEOSSSLNEASSGLF	828		
Db	711	QPVIN-ICQVYINPNGKFKYKOTGIPOGLCVSSILSEFYATLEESSLGRLOESNPE	769		
QY	829	DVLEFMCHH-VIRGRSYVOCGIPGSLSTLSTLSCIGDMENK-L-F--AGIRRD	882		
Db	770	NPVNLRLRLDDYLLITTOENNAVLFEKLIINSRGKFNKKIQTSEPL	822		
QY	883	G---LLRLVDDPLVTPHLTHAKTFLRTLVROGPEGVNLRKIYVNPV	931		
RESULT	7	PRELIMINARY:	PRT:	67 AA.	
ID	035432	AC	035432:		
DT	01-JAN-1998	(TREMBLREL. 05, CREATED)			
DT	01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)			
DT	01-AUG-1998	(TREMBLREL. 07, LAST ANNOTATION UPDATE)			
DE	TELOMERASE CATALYTIC SUBUNIT (FRAGMENT).				
OS	MUS MUSCULUS (MOUSE).				
OC	EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
CC	EUTHERIA; RODENTIA.				
RP	(1)				
RP	SEQUENCE FROM N.A.				
RA	DRISSI R., CLEVELAND J.L.;				
RL	SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL; AF029235; G2605903; -.				
FT	NON-TER	1			
FT	NON-TER	67			
SO	SEQUENCE	67 AA; 8368 MW; E2A06F2B CRC32;			
Query Match	6.0%:	Score 424:	DB 11:	Length 67:	
Best Local Similarity	76.1%:	Pred. No. 1.11e-55:			
Matches 51:	Conservative 10:	Mismatches 6:	Indels 0:	Gaps 0	
Db	1	FFYVVESTFOKNRLFYRKSYWSKLOSIGVQHLERYRLSELQSEVYRHHODTWLAMPIC	60		
QY	582	FFYVETTFQKNRLFYRKSYWSKLOSIGIQHKLRYOLRELSEAEVQHREARPAALTS	641		
Db	61	RLRLETPK	67		
QY	642	RLRLETPK	648		
RESULT	8	PRELIMINARY:	PRT:	884 AA.	
ID	Q06163				

QY 259 RPRRGAPPERTVGGSGMAHPTGRTGSDR-GFCVSPAPPAE-EATLEGALSGTRH 316
 Db 205 APPAGKPGPPPPQGGRRPR 226
 QY 317 SHPSVGRQHHAGPSTSRPRP 338

RESULT 11
 ID 046612 PRELIMINARY; PRT; 276 AA.
 AC 046612;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE IS 1222 GENE ORF-A AND ORF-B.
 OS ENTEROBACTER AGGLOMERANS.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
 OC ENTEROBACTERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KLEBERGER, 1983;
 RX MEDLINE; 95255664.
 RA STEIBL H.D., LEWECKE F.M.;
 RL GENE 156:37-42(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA STEIBL H.D., SIDDAYATTAM D.;
 RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96422758.
 RA STEIBL H.D., SIDDAYATTAM D., KLINGMUELLER W.;
 RL PLASMID 34:223-228(1995).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA STEIBL H.D.;
 RL THESES (1995), UNIVERSITÄT BAYREUTH.
 DR EMBL; X78052; G459248; -;
 DR EMBL; X81893; E258949; -;
 SQ SEQUENCE 276 AA; 31718 MW; 0B54A420 CRC32;
 Query Match 2.0%; Score 141; DB 2; Length 276;
 Best Local Similarity 33.7%; Pred. No. 5.04e-05;
 Matches 30; Conservative 23; Mismatches 30; Indels 6; Gaps 6;
 Db 42 ITTELALERRR-FGYRR-IMQLLRGLVHNKRYRYLYHLSGLGKRRRR-RKGLATERL 98
 QY 585 VTETTFQKNRLFFYRKYSKSLQSIGIRQHLKRV-QLRELSEAEVQRHREARPALTSRL 643
 Db 99 PLL-RPAAPNLWSDDYMDALATGRRIK 126
 QY 644 REIPKPDGLREIYMDYVGA-RTERREK 671
 RESULT 12
 ID 054226 PRELIMINARY; PRT; 373 AA.
 AC 054226;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE POLYKETIDE SYNTHASE (FRAGMENT).
 GN ERYA.
 OS SACCAROPOLYSPORA ERYTHRAEA (STREPTOMYCES ERYTHRAEUS).
 OC PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NREL2338;
 RA SLAH-BET K., DOUMITH M., MICHEL J.M., HAYDOCK S., CORTES J.,
 RA LEADATY P.F., RAYNAL M.C.;
 RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; Y14332; E1228185; -;
 FT NON TER 1
 SQ SEQUENCE 373 AA; 40872 MW; 781BEC6D CRC32;

Query Match 2.0%; Score 140; DB 2; Length 373;
 Best Local Similarity 32.3%; Pred. No. 7.05e-05;
 Matches 40; Conservative 27; Mismatches 46; Indels 11; Gaps 11;

Db 243 PRRAA-NRRRG-AAAPAPAPG-AAHG-HRRG-AARRARRRGGRGRRGGRPG 297
 QY 215 PRRLGGERAMNSVREAGVPLGLPAGARRGGSASRSIPLPKRRRGAPPERTPVG 274
 Db 298 L-RRAAPPR-RGPPRRRRAVGTSRPPRQAG-AGH-R-ARRRRTA-GRGGAARTHAOR 351
 QY 275 QGSAHGRTRGSDRGFCVSPAPPAEATLEGALSGTRHSHPSVGRQHHAGPSTSR 334
 Db 352 PGRP 355
 QY 335 PPRP 338

RESULT 13
 ID 036027 PRELIMINARY; PRT; 574 AA.
 AC 036027;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE WISKOTT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.
 GN WSP1 OR SPACAP10.15C.
 OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
 OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMYCETES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-JS21.
 RA ZANKEL T.C., OW D.W.;
 RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- SIMILARITY: TO YEAST LAS17.
 DR EMBL; AF038575; G2708709; -;
 DR EMBL; Z98980; E339884; -;
 FT DOMAIN 311 317 POLY-PRO.
 FT DOMAIN 337 343 POLY-PRO.
 FT DOMAIN 361 366 POLY-PRO.
 FT VARIANT 248 248 L -> V (IN STRAIN JS21).
 SQ SEQUENCE 574 AA; 59605 MW; 9E84D37C CRC32;

Query Match 2.0%; Score 143; DB 3; Length 574;
 Best Local Similarity 28.8%; Pred. No. 2.56e-05;
 Matches 42; Conservative 41; Mismatches 53; Indels 10; Gaps 7;

Db 305 AANKRRPPPPP-PSRR--NRG-KPPIGNSSNSLPPPPPPRSNAG-SIPLPPQGR 358
 QY 202 AATQARPPPHASGPRRLRGGERAMNSVREAGVPLGLPAGARRGGSASRSIPLPKRR 261
 Db 359 SAPPPPRSAPSTGKRPPLSSRAVSNNP--APPRAIGRANPLPLGNSRSRSTPP 416
 QY 262 RGAAPPERTPVGGSAHAP-GTRGSDRGFCVSPAPPAEATLEGALSGTRHSHPS 320
 Db 417 VPPTP-SLPSAPPSPPSAPSLPM 441
 QY 321 VGRQHHAGPSTSRPRPMDTPCPV 346

RESULT 14
 ID 099076 PRELIMINARY; PRT; 316 AA.
 AC 099076;
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE HOMEBOX PROTEIN (FRAGMENT).
 GN HB9.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

```
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-TONSIL.
RX MEDLINE; 91305125.
RA DEGUCHI Y., KEHRL J.H.;
RL NUCLEIC ACIDS RES. 19:3742-3742(1991).
CC -I- SIMILARITY: STRONG WITH OTHER "ANTP-TYPE" HOMEOBOX PROTEINS.
DR EMBL; X56537; E30256; ALT_SEQ.
DR PIR; S16681; S16681.
DR PFAM; PF00046; homeobox.
KW DNA-BINDING; NUCLEAR PROTEIN; HOMEOBOX; PHOSPHORYLATION.
FT NON_TER 1
FT DNA_BIND 70 129 HOMEOBOX.
FT DNA_BIND 112 121 H-T-H MOTIF.
SQ SEQUENCE 316 AA; 34713 MW; 1531E766 CRC32;

Query Match 1.9%; Score 136; DB 4; Length 316;
Best Local Similarity 30.9%; Pred. No. 2.67e-04;
Matches 21; Conservative 20; Mismatches 24; Indels 3; Gaps 3;

Db 202 RFWSTASDCSVGTGIAIPRGPAT-SPRPSRSPAADRSRPARSAPGPAASFGPGAMTH 260
Oy 223 RAWNHSVREAGV-PLGLPAPGARRGSGSASRLPLPKRPRRG-AAPEPERTPVGOGSWAH 280
Db 261 PARPREQA 268
Oy 281 PERTRPS 288

RESULT 15
ID 062103 PRELIMINARY; PRT; 317 AA.
AC 062103.
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE PROLINE RICH PROTEIN PRECURSOR.
GN PRP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CD-1: TISSUE-LIVER;
RX MEDLINE; 88273214.
RA ANN D.K., SMITH M.K., CARLSON D.M.;
RL J. BIOL. CHEM. 263:10887-10893(1988).
DR EMBL; M23236; G567232; -.
DR MGD; MGI:97773; PRP.
KW SIGNAL.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 317 PROLINE-RICH PROTEIN.
SQ SEQUENCE 317 AA; 31719 MW; 10C84341 CRC32;

Query Match 1.9%; Score 132; DB 11; Length 317;
Best Local Similarity 26.4%; Pred. No. 9.90e-04;
Matches 39; Conservative 38; Mismatches 64; Indels 7; Gaps 6;

Db 37 SSGQRPVPVNGSQGPPPGQPQRP-PQGPQPPPGQPQRP-PQGPQRP-PQGP 94
Oy 202 AATQARPPPHAS-GPRRLGCE-RAWNHSVREAGVPLGLPAPGARRGSGASRSIPLPK 258
Db 95 PPPGPPQPPPGPPPG-PQGPQPPPGPPPGQPQRP-PQGPQPPPGPPPGQPQRP 153
Oy 259 RPRGAAPERTPVGOGSWAHGRTGSDRGFCVSPAR-PAEATSLGALSGTRHS 317
Db 154 GGPQRPQPPPGPPPGQPQRP-PQGPQRP 181
Oy 318 HPSVGRQHHAGPSTSRPPRPMWTPCP 345

Search completed: Fri Dec 18 18:17:04 1998
Job time : 166 secs.
```

This Page Blank (uspto)

Query Match	4.48;	Score 97;	DB 27;	Length 1398,
Best Local Similarity	29.38;	Pred. No. 1.85e+01;		

CC one strand of the DNA substrate is cleaved and a nucleic acid
CC molecule is attached to one or both of the cleaved strands. The

CC receptor designated MEL-1ab, also known as Mel 1-c(beta). The protein differs from previously known receptors by being 65 amino acids shorter and also differs throughout the sequence by 6 amino acids. The protein is encoded by 2 different alleles (T79065-6) which differ in the 3' untranslated region (both sequences encode the same protein), which is thought to affect the half-life of the mRNA. The coding sequence was isolated from cDNA derived from Xenopus skin RNA and amplified using the primers T79067-76. The nucleotide sequence encodes a 7 transmembrane receptor involved in cellular signalling. MEL-1ab has been shown to modulate intracellular cAMP, esp. inhibiting its accumulation induced by an inhibitor of phosphodiesterase, but unlike the MEL-1aa protein, it cannot inhibit adenylyl cyclase activity.

CC Sequence 354 AA;

Query Match 4.1%; Score 90; DB 24; Length 354;

Best Local Similarity 32.1%; Pred. No. 5.64e+01;

Matches 18; Conservative 13; Mismatches 20; Indels 5; Gaps 4;

Db 194 Itvvvvhfvfslsvtfcylfiwvliqvk-rvrgdfkqkltp-talnflmfv 247

Qy 747 LTEVIASIIKQON--TTC-VKRAVAVQKRAHGKAKKSHVSTLTLQPYMKQFV 799

RESULT 8

ID W23652 standard; Protein; 413 AA.

AC W23652;

DT 09-OCT-1997 (first entry)

DE Recombinant squirrel protease inhibitor 55RS.

KW Protease inhibitor; alpha-1 antitrypsin; HP-55; striped squirrel;

KM hypernatation related protein.

OS Tamias asiaticus.

FH Key Location/Qualifiers

FT peptide 1..24

FT protein /label= Signal

FT region /label= Protease_inhibitor_55RS

FT region /label= Reaction_site

PN J09157298-A.

PD 17-JUN-1997.

PF 04-DEC-1995; 315692.

PR 04-DEC-1995; JP-315692.

PA (KONDO/) KONDO J.

DR (KANA-) 2H KANAGAWA KAGAKU GIUTSU ACAD.

DR WPI; 97-369473/34.

DR N-PSDB; T78180.

PT Squirrel derived protease inhibitor has similar structure to alpha-1

PT antitrypsin - shows selective inhibitory activity against

PT protease(s)

PS Claim 4; Page 8-9; 10pp; Japanese.

CC The present sequence represents a recombinant squirrel protease

CC inhibitor protein 55RS that contains a specifically claimed reaction

CC site, see features table. The protease inhibitor 55RS has inhibitory

CC activity against trypsin, elastase or thrombin. The protease inhibitor

CC has a new structure similar to alpha-1-antitrypsin and shows selective

CC inhibitory activity against proteases. The protease inhibitor cDNA was

CC homologous to cDNA encoding a similar squirrel protein designated

CC hibernation-related protein (HP-55).

CC Sequence 413 AA;

Query Match 4.1%; Score 90; DB 23; Length 413;

Best Local Similarity 45.8%; Pred. No. 5.64e+01;

Matches 11; Conservative 6; Mismatches 6; Indels 1; Gaps 1;

Db 295 lknr-qtlksqlyfipkvsigtyd 317

Qy 717 LRVRAQDPPELVFKVDVTGAYD 740

RESULT 9

ID R88409 standard; Protein; 420 AA.

AC R88409;

DT 15-AUG-1996 (first entry)

DE High-affinity melatonin receptor.

KW Melatonin receptor; G-protein-coupled receptor; glycosylation;

KM disulfide bond; ligand binding pocket; phosphorylation;

KW cytochrome-c family heme binding site; melatonin receptor-agonist;

KW melatonin receptor-antagonist; circadian rhythm disorder; jet-lag;

KW day-night cycle disorder; ovulation; reproductive cycle; antibody;

KM animal breeding; puberty; transgenic animal; drug screening.

OS Xenopus laevis.

FH Key Location/Qualifiers

FT domain 1..32

FT modified_site /note= "Extracellular domain"

FT modified_site 5..7

FT modified_site /note= "N-glycosylation site"

FT modified_site 16

FT domain /note= "Protein-kinase-C phosphorylation site"

FT domain 33..57

FT domain /note= "Transmembrane region-I"

FT domain 57..68

FT region /note= "Intracellular loop"

FT region 67..72

FT domain /note= "Conserved melatonin receptor motif"

FT domain 69..93

FT domain /note= "Transmembrane region-II"

FT domain 94..107

FT disulfide_bond /note= "Extracellular loop"

FT domain 105..182

FT domain 108..126

FT region /note= "Transmembrane region-III"

FT region 125..131

FT domain /note= "Region used to construct primer"

FT domain 127..151

FT region /note= "Intracellular loop"

FT region 132..137

FT modified_site /note= "Putative cytochrome-c family heme binding site"

FT domain 137

FT domain /note= "Protein-kinase-C phosphorylation site"

FT domain 152..171

FT misc_difference /note= "Transmembrane region-IV"

FT domain 166

FT domain /note= "Residue which may form ligand binding pocket"

FT domain 172..193

FT domain /note= "Extracellular loop"

FT domain 194..220

FT misc_difference /note= "Transmembrane region-V"

FT domain 204

FT domain /note= "Residue which may form ligand binding pocket"

FT domain 221..243

FT domain /note= "Intracellular loop"

FT domain 244..268

FT region /note= "Transmembrane region-VI"

FT region 252..259

FT misc_difference /note= "Region used to construct primer"

FT domain 238

FT domain /note= "Residue which may form ligand binding pocket"

FT domain 269..279

FT domain /note= "Extracellular loop"

FT domain 280..301

FT region /note= "Transmembrane region-VII"

FT region 296..300

FT domain /note= "Conserved melatonin receptor-1b motif"

FT domain 302..420

FT modified_site /note= "Intracellular domain"

FT modified_site 320

FT modified_site /note= "Protein-kinase-C phosphorylation site"

FT modified_site 328

FT modified_site /note= "Protein-kinase-C phosphorylation site"

FT modified_site 362

FT modified_site /note= "Protein-kinase-C phosphorylation site"

FT modified_site 395

FT modified_site /note= "Protein-kinase-C phosphorylation site"

FT modified_site 418

PN W09535320-A1.

PD 28-DEC-1995.
 PF 07-JUN-1995; U07360.
 PR 17-JUN-1994; US-261857.
 PR 07-OCT-1994; US-319887.
 PR 06-JUN-1995; US-319887.
 PA (MASS-) MASSACHUSETTS GEN HOSPITAL.
 PI Reppert SM;
 DR WPI; 96-058368/06.
 DR N-PDB: T09947.
 PT DNA encoding high affinity melatonin receptor one - used to identify
 PT receptor agonists or antagonists e.g. for regulating circadian
 PT rhythm disorders or reproductive cycles
 PS Claim 5; Fig 1; 115pp; English.
 CC The sequence represents a high-affinity melatonin receptor (mol.wt.
 CC 47,424) from *Xenopus laevis*. The receptor is a membrane protein,
 CC coupled to guanine nucleotide binding proteins (G-proteins), and
 CC has 7 hydrophobic putative transmembrane domains. The N-terminus
 CC has an N-linked glycosylation site typical for G-protein-coupled
 CC receptors, and 2 Cys residues in the 1st 2 extracellular loops may
 CC form a stabilizing disulfide bond. Pro residues in transmembrane
 CC domains IV, V and VI may introduce kinks in the alpha-helices to
 CC form of a ligand binding pocket. Phosphorylation sites in the
 CC C-tail may be involved in receptor regulation. Primers from the
 CC encoding DNA may be used for isolation of sheep, mouse and human
 CC receptor sequences. Receptor fragments which interact with
 CC melatonin, or specific antibodies, may be used as receptor-agonists
 CC or receptor-antagonists. Agonists may be used in therapy of
 CC circadian rhythm disorders such as jet-lag or day-night cycle
 CC disorders, to control ovulation, or in alteration of reproductive
 CC cycles in seasonally breeding animals. Antagonists may be used to
 CC control the initiation or timing of puberty in humans. The
 CC receptor gene may also be expressed in a transgenic animal for use
 CC as a model system to screen agonists and antagonists.
 SQ Sequence 420 AA:

Query Match: 4.1%; Score 90; DB 17; Length 420;
Best Local Similarity 33.9%; Pred.No.5.64e+01;
Matches 19; Conservative 12; Mismatches 20; Indels 5; Gaps 4;

Dd 194 itvvvhtfvlpvsvtfcylrlwlvlgvkh-ryvgdfgkgl-tqtdlrnfltmfv 247
:::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 747 LFEVIASTIKQN--TYC-VRRIVAVOKAANGHVRKAFKSVSLTLDLPIMRQFV 799

RESULT 10
ID R13838 standard; protein; 433 AA.
AC R13838;
DT 20-NOV-1991 (first entry)
DE Fusaric acid resistance protein encoded by fada.
KW Fungl.
OS Pseudomonas cepacia, strain UK-1.
PN EP-444664-A.
PD 04-SEP-1991.
PE 28-FEB-1991; 103006.
PR 28-FEB-1990; JP-045481.
PR 18-FEB-1991; JP-044027.
PA (SUNR) SUNTORI LTD.
PA (DAIK) DAIKIN IND LTD.
PI Shibano Y, Toyoda H, Utsumi R, Obata K;
DR N-PDSB: O13369,
DR N-PDSB: O13369,
PT Fusaric acid resistant genes - derived from fusaric acid
PT decomposing or detoxifying microorganisms e.g. Pseudomonas
PT cepacia.
PS Claim 18; Fig 4; 38pp; English.
CC The sequence was deduced from a clone isolated from a gene library
CC prep'd. from P. cepacia. It is the 1st of five proteins encoded by
CC ORFs fada-fade which are essential for fusaric acid resistance.
CC The protein has a mol. wt. of 47 kD. The genes can be introduced
CC into plants in order to control plant diseases such as wilting
CC disease caused by Fusarium fungl.
See also R13839-R13845.
SO Sequence 433 AA;

Query Match	4.1%	Score 90;	DB 3;	Length 433;
Best Local	Similarity 35.6%;	Pred. 5.64e+01;		
Matches	16;	Conservative	9;	Mismatches 18; Indels 2; Gaps 2;
Db	333	rlprpdrllrgraaiaayahhkkpphsergrlaahaq	lmaal	376
Oy	664	rlrpfpdlrplvmndr-vvgaarffrkkrraerlksrkalfsvl		687

ID RESULT 11
 AC R85478 standard; Protein; 992 AA.
 DE R85478:
 DE 19-JUN-1996 (first entry)
 DE Barnacle adhesion protein.
 KW Barnacle; Megabalanus rosa; adhesion protein; adhesive; water
 OS Megabalanus rosa.
 PN J07265081.A.
 PD 17-OCT-1995.
 PF 31-MAR-1994; 064036.
 PR 31-MAR-1994; JP-064036.
 PA (KAIY-) KAIYO BIOTECHNOLOGY KENKYUSHO KK.
 DR WPI; 95-388689/50.
 DR N-PSDB; T057722.
 PT Barnacle adhesion protein gene - for production of adhesive protein
 PS Claim 1; Page 7; 7pp; Japanese.
 CC This is the amino acid sequence of the barnacle (Megabalanus rosa)
 CC adhesion protein. The corresp. gene was isolated from a cDNA library
 CC using the probe T05723 which is derived from amino acid sequencing
 CC of the purified protein. The gene is used to produce an adhesive
 CC protein which can be used in water.
 CC Sequence 992 AA;

	Query Match	4.1%;	Score 90;	DB 16;	Length 992;
	Best Local Similarity	24.2%;	Pred. No. 5,64e+01;		
	Matches	23;	Conservative	28;	Mismatches 37; Indels 7; Gaps 7.
Db	849	irsiv-ahctlllrlarysitcncqyhltkyyngipinalgaafaynlly-clrvpakra	906		
		: : : : : : : : : :			
QY	751	IASIIKPNNTCYVR-KRAVVOGKAHGHVKAFKS-HYSTLIDLQY-MRQFVAHLDETSP	807		
Db	907	aissvgy-qsa-maelymhmrlgmfpssdvri	939		
		: : : : :			
QY	808	LRDAAVIEQSSLENAASSGLEFDVFRLRMCHHAAYRI	842		

RESULT	12	standard: Protein: 1784 AA.
ID	R94427	
AC	R94427;	
DT	11-JUN-1996	(first entry)
DE	Felyv F6A provirus clone 61E encoded nucleocapsid protein.	
KW	Felyv; retrovirius; vaccine; AIDS; disease model; immunodeficiency;	
KW	viraemia; leukaemia; therapy; nucleocapsid.	
OS	Feline immunodeficiency virus subtype A.	
Key	Location/Qualifiers	
FT	misc.difference	581
FT		
FT	EP-699758-AI.	/note="codon 581 in encoding sequence is TAG"
PD	06-MAR-1996.	
PF	12-DEC-1989;	100871.
PR	13-DEC-1988;	US-284139.
PA	(COLS) UNIV COLORADO STATE RES FOUND.	
PA	(HARD) UNIV HARVARD.	
PI	Hoover EA, Mullins JT;	
DR	WPI: 96-180826/19.	
DR	N-PSDB: T13265.	
FT	Inactivated Felyv-A sub-type isolates for use in disease models and	
FT	vaccines - can be used to study prophylaxis and therapy of related	
FT	immuno-deficiencies in other species, e.g. human.	
PS	Disclosure: Fig 1: 22pp; English.	
CC	The nucleocapsid protein (R94427) and envelope protein (R94428)	
CC	sequences of feline leukemia virus provirus F6A clone 61E were	
CC	deduced from open reading frames identified in the proviral DNA	

PI Young D, Zhang Y;
DR WPL: 93-368812/46.

Query Match	3.98;	Score 86;	DB 8;	Length 1702;
Best Local Similarity	28.68;	Pred. No. 1.05e+02;		
Matches	22;	Conservative	19;	Mismatches 28; Indels 8; Gaps 7;

THIS PAGE BLANK (USPTO)

MPSRCH_PP protein - protein database search, using Smith-Waterman algorithm
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:36:16 1998; Maspar time 12.67 Seconds
811.235 Million cell updates/sec
Tabular output not generated.

Title: >US-08-951-733-14
Description: (640-940) from US08951733.pep (2 of 2)
Perfect Score: 2214
Sequence: 1 TSRLRFIPKPGRLRPVAMD.....NLRKVVNFPVEDEALGGTA 301

Scoring table:
PAM 150
Gap 11

Searched: 107076 seqs, 34141958 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: p1r56
1:p1r1 2:p1r2 3:p1r3 4:p1r4

Statistics: Mean 47.293; Variance 102.661; scale 0.461

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	148	6.7	884	2	S53396	1.84e-06
2	108	4.9	94	2	H70127	hypothetical protein
3	108	4.9	366	2	A64591	hypothetical protein
4	103	4.7	189	2	S49743	GTP-binding protein -
5	105	4.7	891	2	G64875	hypothetical protein
6	103	4.7	891	2	B48642	aconitate hydratase (
7	102	4.6	205	2	S69504	aconitate hydratase (
8	102	4.6	318	2	S58503	reverse transcriptase
9	101	4.6	551	1	ORECS	serine transcarboxylase
10	101	4.6	1114	2	SS1470	hypothetical protein
11	101	4.6	3411	1	GNMVP	hypothetical protein -
12	101	4.6	3411	1	GNMVP	genome polyploid protein -
13	99	4.5	372	2	A64462	8-amino-7-oxononanoate
14	99	4.5	446	2	S32871	regulatory protein mo
15	100	4.5	557	2	S25023	neurofilament-like pr
16	99	4.5	1176	2	S66771	hypothetical protein
17	98	4.4	586	2	S19381	hypothetical protein
18	97	4.4	1238	2	S68700	hypothetical protein
19	96	4.3	308	2	E69771	HIPP beta-like tyrosi
20	96	4.3	318	1	NMBE78	ABC transporter (Atp-
21	96	4.3	363	2	A48338	UL18 protein - human
22	96	4.3	417	2	D64302	hypothetical protein
23	96	4.3	789	1	QXB372	L-asparaginase I - Me

24	96	4.3	854	2	S17995	gene COX1 intron 1 pr
25 <td>92<td>4.2<td>212<td>2<td>S73466<th>probable DNA primase</th></td></td></td></td></td>	92 <td>4.2<td>212<td>2<td>S73466<th>probable DNA primase</th></td></td></td></td>	4.2 <td>212<td>2<td>S73466<th>probable DNA primase</th></td></td></td>	212 <td>2<td>S73466<th>probable DNA primase</th></td></td>	2 <td>S73466<th>probable DNA primase</th></td>	S73466 <th>probable DNA primase</th>	probable DNA primase
26 <td>93<td>4.2<td>275<td>2<td>F65076<th>hypothetical protein</th></td></td></td></td></td>	93 <td>4.2<td>275<td>2<td>F65076<th>hypothetical protein</th></td></td></td></td>	4.2 <td>275<td>2<td>F65076<th>hypothetical protein</th></td></td></td>	275 <td>2<td>F65076<th>hypothetical protein</th></td></td>	2 <td>F65076<th>hypothetical protein</th></td>	F65076 <th>hypothetical protein</th>	hypothetical protein
27 <td>93<td>4.2<td>363<td>2<td>I51247<th>aldolase C - African</th></td></td></td></td></td>	93 <td>4.2<td>363<td>2<td>I51247<th>aldolase C - African</th></td></td></td></td>	4.2 <td>363<td>2<td>I51247<th>aldolase C - African</th></td></td></td>	363 <td>2<td>I51247<th>aldolase C - African</th></td></td>	2 <td>I51247<th>aldolase C - African</th></td>	I51247 <th>aldolase C - African</th>	aldolase C - African
28 <td>93<td>4.2<td>364<td>2<td>S45346<th>fructose-bisphosphate</th></td></td></td></td></td>	93 <td>4.2<td>364<td>2<td>S45346<th>fructose-bisphosphate</th></td></td></td></td>	4.2 <td>364<td>2<td>S45346<th>fructose-bisphosphate</th></td></td></td>	364 <td>2<td>S45346<th>fructose-bisphosphate</th></td></td>	2 <td>S45346<th>fructose-bisphosphate</th></td>	S45346 <th>fructose-bisphosphate</th>	fructose-bisphosphate
29 <td>92<td>4.2<td>442<td>2<td>A69873<th>phosphate starvation</th></td></td></td></td></td>	92 <td>4.2<td>442<td>2<td>A69873<th>phosphate starvation</th></td></td></td></td>	4.2 <td>442<td>2<td>A69873<th>phosphate starvation</th></td></td></td>	442 <td>2<td>A69873<th>phosphate starvation</th></td></td>	2 <td>A69873<th>phosphate starvation</th></td>	A69873 <th>phosphate starvation</th>	phosphate starvation
30 <td>93<td>4.2<td>527<td>2<td>A53467<th>protein kinase SNF1 h</th></td></td></td></td></td>	93 <td>4.2<td>527<td>2<td>A53467<th>protein kinase SNF1 h</th></td></td></td></td>	4.2 <td>527<td>2<td>A53467<th>protein kinase SNF1 h</th></td></td></td>	527 <td>2<td>A53467<th>protein kinase SNF1 h</th></td></td>	2 <td>A53467<th>protein kinase SNF1 h</th></td>	A53467 <th>protein kinase SNF1 h</th>	protein kinase SNF1 h
31 <td>93<td>4.2<td>581<td>2<td>G70189<th>threonine--cRNA ligas</th></td></td></td></td></td>	93 <td>4.2<td>581<td>2<td>G70189<th>threonine--cRNA ligas</th></td></td></td></td>	4.2 <td>581<td>2<td>G70189<th>threonine--cRNA ligas</th></td></td></td>	581 <td>2<td>G70189<th>threonine--cRNA ligas</th></td></td>	2 <td>G70189<th>threonine--cRNA ligas</th></td>	G70189 <th>threonine--cRNA ligas</th>	threonine--cRNA ligas
32 <td>92<td>4.2<td>647<td>2<td>E64170<th>hypothetical protein</th></td></td></td></td></td>	92 <td>4.2<td>647<td>2<td>E64170<th>hypothetical protein</th></td></td></td></td>	4.2 <td>647<td>2<td>E64170<th>hypothetical protein</th></td></td></td>	647 <td>2<td>E64170<th>hypothetical protein</th></td></td>	2 <td>E64170<th>hypothetical protein</th></td>	E64170 <th>hypothetical protein</th>	hypothetical protein
33 <td>94<td>4.2<td>781<td>2<td>E64222<th>topoisomerase IV chai</th></td></td></td></td></td>	94 <td>4.2<td>781<td>2<td>E64222<th>topoisomerase IV chai</th></td></td></td></td>	4.2 <td>781<td>2<td>E64222<th>topoisomerase IV chai</th></td></td></td>	781 <td>2<td>E64222<th>topoisomerase IV chai</th></td></td>	2 <td>E64222<th>topoisomerase IV chai</th></td>	E64222 <th>topoisomerase IV chai</th>	topoisomerase IV chai
34 <td>93<td>4.2<td>1766<td>2<td>S03701<th>lipik protein - pea ea</th></td></td></td></td></td>	93 <td>4.2<td>1766<td>2<td>S03701<th>lipik protein - pea ea</th></td></td></td></td>	4.2 <td>1766<td>2<td>S03701<th>lipik protein - pea ea</th></td></td></td>	1766 <td>2<td>S03701<th>lipik protein - pea ea</th></td></td>	2 <td>S03701<th>lipik protein - pea ea</th></td>	S03701 <th>lipik protein - pea ea</th>	lipik protein - pea ea
35 <td>91<td>4.1<td>164<td>1<td>CEXCA<th>C-phycocyanin alpha</th></td></td></td></td></td>	91 <td>4.1<td>164<td>1<td>CEXCA<th>C-phycocyanin alpha</th></td></td></td></td>	4.1 <td>164<td>1<td>CEXCA<th>C-phycocyanin alpha</th></td></td></td>	164 <td>1<td>CEXCA<th>C-phycocyanin alpha</th></td></td>	1 <td>CEXCA<th>C-phycocyanin alpha</th></td>	CEXCA <th>C-phycocyanin alpha</th>	C-phycocyanin alpha
36 <td>91<td>4.1<td>352<td>2<td>F69117<th>conserved hypothetical</th></td></td></td></td></td>	91 <td>4.1<td>352<td>2<td>F69117<th>conserved hypothetical</th></td></td></td></td>	4.1 <td>352<td>2<td>F69117<th>conserved hypothetical</th></td></td></td>	352 <td>2<td>F69117<th>conserved hypothetical</th></td></td>	2 <td>F69117<th>conserved hypothetical</th></td>	F69117 <th>conserved hypothetical</th>	conserved hypothetical
37 <td>91<td>4.1<td>357<td>2<td>ADSPAC<th>fructose-bisphosphate</th></td></td></td></td></td>	91 <td>4.1<td>357<td>2<td>ADSPAC<th>fructose-bisphosphate</th></td></td></td></td>	4.1 <td>357<td>2<td>ADSPAC<th>fructose-bisphosphate</th></td></td></td>	357 <td>2<td>ADSPAC<th>fructose-bisphosphate</th></td></td>	2 <td>ADSPAC<th>fructose-bisphosphate</th></td>	ADSPAC <th>fructose-bisphosphate</th>	fructose-bisphosphate
38 <td>91<td>4.1<td>366<td>2<td>S56487<th>hypothetical 40.4k pr</th></td></td></td></td></td>	91 <td>4.1<td>366<td>2<td>S56487<th>hypothetical 40.4k pr</th></td></td></td></td>	4.1 <td>366<td>2<td>S56487<th>hypothetical 40.4k pr</th></td></td></td>	366 <td>2<td>S56487<th>hypothetical 40.4k pr</th></td></td>	2 <td>S56487<th>hypothetical 40.4k pr</th></td>	S56487 <th>hypothetical 40.4k pr</th>	hypothetical 40.4k pr
39 <td>91<td>4.1<td>609<td>2<td>F64045<th>exonuclease ABC, cha</th></td></td></td></td></td>	91 <td>4.1<td>609<td>2<td>F64045<th>exonuclease ABC, cha</th></td></td></td></td>	4.1 <td>609<td>2<td>F64045<th>exonuclease ABC, cha</th></td></td></td>	609 <td>2<td>F64045<th>exonuclease ABC, cha</th></td></td>	2 <td>F64045<th>exonuclease ABC, cha</th></td>	F64045 <th>exonuclease ABC, cha</th>	exonuclease ABC, cha
40 <td>91<td>4.1<td>709<td>2<td>C28821<th>1-phosphatidylinosito</th></td></td></td></td></td>	91 <td>4.1<td>709<td>2<td>C28821<th>1-phosphatidylinosito</th></td></td></td></td>	4.1 <td>709<td>2<td>C28821<th>1-phosphatidylinosito</th></td></td></td>	709 <td>2<td>C28821<th>1-phosphatidylinosito</th></td></td>	2 <td>C28821<th>1-phosphatidylinosito</th></td>	C28821 <th>1-phosphatidylinosito</th>	1-phosphatidylinosito
41 <td>91<td>4.1<td>780<td>1<td>A31070<th>6-phosphofructokinase</th></td></td></td></td></td>	91 <td>4.1<td>780<td>1<td>A31070<th>6-phosphofructokinase</th></td></td></td></td>	4.1 <td>780<td>1<td>A31070<th>6-phosphofructokinase</th></td></td></td>	780 <td>1<td>A31070<th>6-phosphofructokinase</th></td></td>	1 <td>A31070<th>6-phosphofructokinase</th></td>	A31070 <th>6-phosphofructokinase</th>	6-phosphofructokinase
42 <td>91<td>4.1<td>989<td>2<td>I56333<th>apolipoprotein B - ra</th></td></td></td></td></td>	91 <td>4.1<td>989<td>2<td>I56333<th>apolipoprotein B - ra</th></td></td></td></td>	4.1 <td>989<td>2<td>I56333<th>apolipoprotein B - ra</th></td></td></td>	989 <td>2<td>I56333<th>apolipoprotein B - ra</th></td></td>	2 <td>I56333<th>apolipoprotein B - ra</th></td>	I56333 <th>apolipoprotein B - ra</th>	apolipoprotein B - ra
43 <td>91<td>4.1<td>1221<td>1<td>HJNVAV<th>helicase (EC 3.6.1.-)</th></td></td></td></td></td>	91 <td>4.1<td>1221<td>1<td>HJNVAV<th>helicase (EC 3.6.1.-)</th></td></td></td></td>	4.1 <td>1221<td>1<td>HJNVAV<th>helicase (EC 3.6.1.-)</th></td></td></td>	1221 <td>1<td>HJNVAV<th>helicase (EC 3.6.1.-)</th></td></td>	1 <td>HJNVAV<th>helicase (EC 3.6.1.-)</th></td>	HJNVAV <th>helicase (EC 3.6.1.-)</th>	helicase (EC 3.6.1.-)
44 <td>91<td>4.1<td>1659<td>1<td>QXB372<th>protein kinase GCN2 (</th></td></td></td></td></td>	91 <td>4.1<td>1659<td>1<td>QXB372<th>protein kinase GCN2 (</th></td></td></td></td>	4.1 <td>1659<td>1<td>QXB372<th>protein kinase GCN2 (</th></td></td></td>	1659 <td>1<td>QXB372<th>protein kinase GCN2 (</th></td></td>	1 <td>QXB372<th>protein kinase GCN2 (</th></td>	QXB372 <th>protein kinase GCN2 (</th>	protein kinase GCN2 (
45 <td>91<td>4.1<td>1797<td>2<td>F69195<th>cell surface glycopro</th></td></td></td></td></td>	91 <td>4.1<td>1797<td>2<td>F69195<th>cell surface glycopro</th></td></td></td></td>	4.1 <td>1797<td>2<td>F69195<th>cell surface glycopro</th></td></td></td>	1797 <td>2<td>F69195<th>cell surface glycopro</th></td></td>	2 <td>F69195<th>cell surface glycopro</th></td>	F69195 <th>cell surface glycopro</th>	cell surface glycopro

RESULT 1
ENTRY S53396 #type complete
TITLE hypothetical protein YLR318w - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES hypothetical protein L8543.12
ORGANISM formal name Saccharomyces cerevisiae
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change 06-Feb-1998

ACCESSIONS S53396
REFERENCE S53390

#authors Du, Z.
#submission submitted to the EMBL Data Library, February 1995
#description The sequence of S. cerevisiae cosmid 8543.
#accession S53396

##molecule_type DNA
##residues 1-884 ##label DU2
##cross-references EMBL:U20618; NID:q2258165; PID:g662136; MIPS:YLR318w
##experimental_source strain S288C (AB972)

GENETICS
#gene SGD:EST2
#map_position 12R
#cross-references SGD:S0004310; MIPS:YLR318w

SUMMARY
#map_position 12R #length 884 #molecular_weight 102662 #checksum 7604

Query Match 6.7%, Score 148, DB 2; Length 884;
Best Local Similarity 26.2%; Pred. No. 1.84e-06;
Matches 68; Conservative 67; Mismatches 93; Indels 32; Gaps 25;

Db	481	ILEYLRKRPSTF-TKLYSPQIADRI-KEFKORLLKFNUNVDELKFMKFDVSCYDSI	538
Qy	686	VLNERRARPELGLASVGLDDI-HRMKRTFVLRV-BA-QDPPELFLVYDVYGAIDTI	742
Db	539	PRMCKMR-ILKDALKNNGFVRISQY-FFNTNT-G-VLKLFNV-VNARSVPKPY-EDYID	592
Qy	743	PQ-DRLTEVIASIRKPNQTCVR-RVAVVOKAANGHAKKAFKSVSLTLDLPFMRQFVA	800
Db	593	NVR-TVHLSNODVIN--VV-E---M-EIF-K-T---ALWVEDKCIREDGLFQSSLS	637
Qy	801	HLQETSLRLDAVVLEQSSSINEASSGFLDFELRMCHHAAVIRKSSVOCGIPGSGILS	860
Db	638	APIYDLYDDILEYSEFKASPSQDTLILKLADFLISTDQOVIVNIKILAMGFGQYVN	697
Qy	861	TLLCSLCYGDH-E-NKLF-AGIRRDGLRLKYDDFLVTLHLTHAKTFLRLVNGVPEYG	917
Db	698	AKANRDKILAVSOSDDDTV 717	
Qy	918	CVVNLKRTV-VNFPVEDEAL 936	

```
RESULT 2
ENTRY 2
TITLE H70127 #type complete
#formal_name Borrelia burgdorferi #common_name Lyme disease
ORGANISM spirochete
DATE 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change
13-Feb-1998

ACCESSIONS
REFERENCE H70127
#authors Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.;
Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Uterback, T.; Matthey, L.; McDonald, L.; Artlich, P.;
Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.

#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochaete, Borrelia
#cross-references M01D:98065943
#accession H70127
#status preliminary: nucleic acid sequence not shown;
translation not shown

#molecule_type DNA
#residues 1-94 #label KLE
#cross-references GB:AE000783; TIGR:BB0224
#experimental_source strain B31

SUMMARY #length 94 #molecular_weight 10659 #checksum 9000

Query Match 4.9%; Score 108; DB 2; Length 94;
Best Local Similarity 26.8%; Pred. No. 3.12e-01;
Matches 19; Conservative 19; Mismatches 30; Indels 3; Gaps 3;

Db 12 LFLSFIISICRKGIVYKNGCINKGISEKEKILLASVSCNHYDINSINDGKANN 71
QY 827 LRFVVF-LRFVCH-HAVAKIRKSTVQCQGIPOGSLT-STLLCSLCYGDMEKRLFRAGIRRG 883
DB 72 LNLSTKIKAPLL 82
QY 884 LLRLVDDFL 894

RESULT 3
ENTRY 3
TITLE A64591 #type complete
#formal_name Helicobacter pylori (strain 26695)
ORGANISM 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change
DATE 10-Oct-1997

ACCESSIONS
REFERENCE A64591
#authors Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.;
Sutton, G.G.; Fleischmann, R.D.; Ketchum, K.A.; Klenk,
H.P.; Gill, S.; Dougherty, B.A.; Nelson, K.; Quackenbush,
J.; Zhou, L.; Kirkness, E.F.; Peterson, S.; Loftus, B.;
Richardson, D.; Dodson, R.; Khalaf, H.G.; Glodek, A.;
McKenney, K.; Fitzgerald, L.M.; Lee, N.; Adams, M.D.;
Hickey, E.K.; Berg, D.E.; Gocayne, J.D.; Uterback, T.R.;
Peterson, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.;
Fujii, C.; Bowman, C.; Matthey, L.; Mallin, E.; Hayes,
W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
C.M.; Venter, J.C.
#journal Nature (1997) 388:539-547
#title The complete genome sequence of the gastric pathogen
#cross-references M01D:97394467
#accession A64591
#status preliminary: nucleic acid sequence not shown;
translation not shown

#molecule_type DNA

##residues 1-366 ##label TOM
##cross-references GB:AE000571; GB:AE000511; NID:g2313686; PID:g2313689;
TIGR:HP0569

SUMMARY #length 366 #molecular_weight 40574 #checksum 4119

Query Match 4.9%; Score 108; DB 2; Length 366;
Best Local Similarity 25.8%; Pred. No. 3.12e-01;
Matches 16; Conservative 21; Mismatches 22; Indels 3; Gaps 3;

Db 3 LSYGIVGLPNVGS-STFNALTKTQNAQSANTPCTIEPNKAIYVNPDRRL-DALAQIVK 60
QY 698 LGSYVGLDLDIHAWMTFVLVRADPPPELY-FVAVDTGAYDRIPODLRTLEVIASIIK 756
DB 61 PE 62
QY 757 PQ 758

RESULT 4
ENTRY 4
TITLE S49743 #type complete
#formal_name Saccharomyces cerevisiae
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 13-Jan-1995 #sequence_revision 10-Feb-1995 #text_change
21-Nov-1997

ACCESSIONS
REFERENCE S49743
#authors Badcock, K.; Churcher, C.
#submission submitted to the EMBL Data Library, November 1994
#accession S49743
#molecule_type DNA
#residues 1-189 ##label BAD
#cross-references EMBL:246659; NID:g5755680; PID:g5755683; MIPS:YML036w

GENETICS #map_position 13L
SUMMARY #length 189 #molecular_weight 21771 #checksum 9333

Query Match 4.7%; Score 103; DB 2; Length 189;
Best Local Similarity 26.1%; Pred. No. 1.20e-00;
Matches 23; Conservative 25; Mismatches 35; Indels 5; Gaps 5;

Db 50 GROMYSAYIKTLEIV-KYKMKRTNINSECVLCLSPTSNISDAF-LKFGIKDSSOLICL 107
QY 673 AERTLSRV-KALFSVINTRRARPGLGASVGLDDIHAWMTFVLVRADPPPELYTV 731
DB 108 KFH-TNT-DDVDKEQRLTIMISIVKGE 133
QY 732 KVDVTGAYDRIPODLRTLEVIASIIKPN 759

RESULT 5
ENTRY 5
TITLE G64875 #type complete
#formal_name Escherichia coli
ORGANISM #formal_name Escherichia coli
DATE 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change
05-Dec-1997

ACCESSIONS
REFERENCE G64875; S22375; A49756
#authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.;
Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.;
Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.;
Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao,
Y.
#journal Science (1997) 277:1453-1462
#title The complete genome sequence of Escherichia coli K-12.
#cross-references M01D:97426617
#accession G64875
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-891 ##label BLAT
#cross-references GB:AE000225; GB:U00096; NID:g1787523; PID:g1787531;
UMCP:b1276
#experimental_source strain K-12, substrain MG1655
```

REFERENCE	S22374
#authors	Prodromou, C.; Artymkiv, P.J.; Guest, J.R.
#journal	Eur. J. Biochem. (1992) 204:599-609
#title	The acontase of Escherichia coli. Nucleotide sequence of the acontase gene and amino acid sequence similarity with mitochondrial acontases, the iron-responsive-element-binding protein and isopropylmalate isomerases.
#accession	S22375
#molecule_type	DNA
#residues	1-521, /G/, 523-891 #label PRD
#cross-references	EMBL:X60293; NID:g40894; PID:g40896
REFERENCE	A49756
#authors	Prodromou, C.; Haynes, M.J.; Guest, J.R.
#journal	J. Gen. Microbiol. (1991) 137:2505-2515
#title	The acontase of Escherichia coli: purification of the enzyme and molecular cloning and map location of the gene (acn).
#accession	A49756
#molecule_type	protein
#residues	'X', 3-15, 'X', 17, 'X', 19 #label PR2
GENETICS	
#gene	acnA; acn
FUNCTION	
#description	reversibly catalyzes the hydration of cis-acontate to citrate and also to isocitrate
#pathway	tricarballic acid cycle
KEYWORDS	4Fe-4S; carbon-oxygen lyase; hydro-lyase; iron-sulfur protein; tricarboxylic acid cycle
FEATURE	
2-891	#product acontate hydratase #status predicted #label MAR\
128,129,181,212,	#active-site Asp, His, His, Asp, His, Glu, Ser #status predicted\
214,309,776	#binding site 4Fe-4S cluster (Cys) (covalent) #status predicted
435,501,504	#length 891 #molecular-weight 97676 #checksum 2334
SUMMARY	
Query Match	4.7%; Score 105; DB 2; Length 891;
Best Local Similarity	21.9%; Pred. No. 7, 05e-01;
Matches 14; Conservative	22; Mismatches 25; Indels 3; Gaps 3;
Db	755 MAYKODPLVIAGKEYSSGR-DMAKRPLLG-IRYVIE-SFERIHRNSLTGMGT 811 ::: :: : : : : : : :: : : :
QY	643 LRFIPKPOLRIIVNDYVGARTFRERKAERLTSRVKALFSLVNERARRRGGLGASV 702
Db	812 LPLE 815
QY	: :
	703 LGLD 706
RESULT	6
ENTRY	B48642
TITLE	#type complete
ORGANISM	aconitate hydratase (EC 4.2.1.3) - Legionella pneumophila
DATE	#formal_name Legionella pneumophila 03-May-1994 #sequence_revision 03-May-1994 #text-change 09-Sep-1997
ACCESSIONS	B48642
REFERENCE	A48642
#authors	Mengaud, J.M.; Horvitz, M.A.
#journal	J. Bacteriol. (1993) 175:5665-5676
#title	The major iron-containing protein of Legionella pneumophila is an aconitase homologous with the human iron-responsive element-binding protein.
#accession	B48642
#status	Preliminary
#molecule_type	DNA
#residues	1-891 #label MEN
#cross-references	GB:L22081; NID:g348943; PID:g348945
KEYWORDS	carbon-oxygen lyase; hydro-lyase
SUMMARY	#length 891 #molecular-weight 99208 #checksum 4922
Query Match	4.7%; Score 103; DB 2; Length 891;

	Best Local Similarity	25.0%	Pred. No.	1,20e+00;	Mismatches	23;	Indels	3;	Gaps	3;
	Matches	16;	Conservative	22;	Mismatches	23;	Indels	3;	Gaps	3;
Dd	755	MRYDENODVLIAGKEVGTSST-DWAAGTNILG-VKAVITE-SFERIHSNLIOMGI	811							
Oy	643	LRFIKPGLPRIVNMDDYVGARFFREKRERLTSVKALFSVLNERARRRGLLGASV	702							
Dd	812	LPLO	815							
Oy	703	LGLD	706							
RESULT	7									
ENTRY		S69504	#type complete							
TITLE		hypothetical protein I - phase HPI								
ORGANISM		#formal_name phase HPI								
DATE		06-Feb-1996 #sequence_revision 13-Mar-1997 #text_change 31-Oct-1997								
ACCESSIONS		S69504: S72330								
REFERENCE		S69503								
#authors		Espósito, D.; Fitzmaurice, W.P.; Benjamin, R.C.; Goodman, S.D.; Waldman, A.S.; Scocca, J.J.								
#journal		Nucleic Acids Res. (1996) 24:2360-2368								
#title		The complete nucleotide sequence of bacteriophage HPI DNA.								
#accession		S69504								
#status		preliminary; nucleic acid sequence not shown; translation not shown								
##molecule_type		DNA								
##residues		1-205 #label ESP								
##cross-references		EMBL:U24159; NID:g1046235; PID:g459176								
#note		the nucleotide sequence was submitted to the EMBL Data Library, April 1995								
REFERENCE		S72329								
#authors		Espósito, D.; Scocca, J.J.								
#journal		Mol. Microbiol. (1994) 13:685-695								
#title		Identification of an Hpi phage protein required for site-specific excision.								
#accession		S72330								
#status		preliminary; nucleic acid sequence not shown; translation not shown								
##molecule_type		DNA								
##residues		1-205 #label ES2								
##cross-references		GB:U24155; EMBL:U06847; NID:g1046235; PID:g459176								
#note		the nucleotide sequence was submitted to the EMBL Data Library, February 1994								
SUMMARY		#length 205 #molecular_weight 23260 #checksum 7617								
Query Match		4.6%; Score 102; DB 2; Length 205;								
Best Local Similarity		22.1%; Pred. NO. 1.5E+00;								
Matches		33; Conservative								
		49; Mismatches 50; Indels 17; Gaps 17;								
Dd	52	SLSTIIICFCFAYGMSWKOTPD-LSKY-AK-NIIDFD-TN-AVLPSKKQSIDET	106							
Oy	651	GIRPLVNMDY-VGARFFREKRERLTSVKALFSVLNERARRRGLLGASV-LGDOI	708							
Dd	107	-KEYNALDKIFWIDEHLSLTLYFKSN-KOEITDITGLADINSATSLIEKHARYDE	164							
Oy	709	HRAWTEFLVAR-RADPP-PELY-FVKVDYTGAYDIRPODLTEVI-A-SIIRPMTYC-	762							
Dd	165	VGRYOVLINNGY-KNTY-PNTKTLEL	191							
Oy	763	VRRYAIVOKAAHGVRKAFAKSHVTIDL	791							
RESULT	8									
ENTRY		S58503	#type fragment							
TITLE		reverse transcriptase like protein 1, intron-encoded - brown alga (Pyrallela littoralis) (fragment)								
ORGANISM		#formal_name Pyrallela littoralis								
DATE		29-Nov-1995 #sequence_revision 01-Mar-1996 #text_change 01-Mar-1996								
ACCESSIONS		S58503								
REFERENCE		S58503								

#authors Fontaine, J.M.; Rousvoal, S.; Leblanc, C.; Kloareg, B.;
 #journal J. Mol. Biol. (1995) 251:378-389
 #title The mitochondrial LSU rRNA of the brown alga *Pyralella littoralis* reveals alpha-proteobacterial features and is split by four group IIB introns with an atypical phylogeny.

#accession S58503
 #status preliminary
 #molecule_type DNA
 #residues 1-318 #label FON
 SUMMARY #length 318 #checksum 7326

Query Match 4.6%; Score 102; DB 2; Length 318;
 Best Local Similarity 46.7%; Pred. NO.1.57e+00;
 Matches 14; Conservative 7; Mismatches 8; Indels 1; Gaps 1;

Db 126 SKSY-PSKRGTVGGVLSPLLCNMNLINGLEN 154
 QY 844 GKSYVOCGIRPGSITSLTCLCYGDMEN 873

RESULT 9
 ENTRY ORCS #type complete
 TITLE serine chemoreceptor protein - Escherichia coli
 ALTERNATE_NAMES methyl-accepting chemotaxis protein I (MCP-I)
 ORGANISM #formal_name Escherichia coli
 DATE 14-Nov-1983 #sequence_revision 13-Feb-1998 #text_change 20-Mar-1998
 E65250; S56581; A03442; S30281
 A64720
 #accessions E65250; S56581; A03442; S30281
 REFERENCE
 #authors Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor, J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.

#journal Science (1997) 277:1453-1462
 #title The complete genome sequence of Escherichia coli K-12.
 #cross-references MIMD:9742617
 #accession E65250
 #status preliminary; nucleic acid sequence not shown; translation not shown

#molecule_type DNA
 #residues 1-551 #label BLAT
 #cross-references GB:A000506; GB:U00096; NID:92367377; PID:92367378; UMGCP:D4355

REFERENCE
 #experimental_source strain K-12, substrain MG1655
 #journal S56314
 #authors Blattner, F.R.; Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Nucleic Acids Res. (1995) 23:2105-2119
 #title Analysis of the Escherichia coli genome VI: DNA sequence of the region from 92.8 through 100 minutes.
 #cross-references MIMD:9534362
 #accession S56581
 #status preliminary; nucleic acid sequence not shown; translation not shown

#molecule_type DNA
 #residues 1-503 'X', 505-551 #label BUR
 #cross-references EMBL:U14003; NID:91263172; PID:9537197
 #note the nucleotide sequence was submitted to the EMBL Data Library, August 1994

REFERENCE
 #journal A03442
 #authors Boyd, A.; Kendall, K.; Simon, M.I.
 #journal Nature (1983) 301:623-626
 #title Structure of the serine chemoreceptor in Escherichia coli.
 #cross-references MIMD:83141753
 #accession A03442
 #molecule_type DNA
 #residues 1-147 'S'RSSLSIRPRDIRN', 162-536 #label BOY
 #cross-references GB:V00373; GB:J01718; NID:943217; PID:943218
 REFERENCE
 #journal S30279
 #authors Roper, D.I.; Pawcett, T.; Cooper, R.A.
 #journal Mol. Gen. Genet. (1993) 237:241-250

#title The Escherichia coli C homeoproteocatechuate degradative operon: hpc gene order, direction of transcription and control of expression.

#accession S30281
 #molecule_type DNA
 #residues 1-77 #label ROP
 #cross-references GB:S56952; NID:9298654; PID:9298656
 COMMENT This protein responds to changes in serine concentration in the environment, transduces a signal from the outside to the inside of the cell, and facilitates sensory adaptation through various levels of methylation.

GENETICS
 #gene tsr
 #map_position 99 min
 CLASSIFICATION #superfamily methyl-accepting chemotaxis protein
 KEYWORDS chemotaxis; methylated amino acid; sensory transduction; transmembrane protein

FEATURE
 297,311 #modified_site glutamate methyl ester (Glu) (by cheB-dependent deamidation and methylation) #status experimental
 304,493 #modified_site glutamate methyl ester (Glu) #status experimental
 503 #modified_site glutamate methyl ester (Glu) #status predicted

SUMMARY #length 551 #molecular_weight 59442 #checksum 1281

Query Match 4.6%; Score 101; DB 1; Length 551;
 Best Local Similarity 22.4%; Pred. NO.2.04e+00;
 Matches 22; Conservative 28; Mismatches 44; Indels 4; Gaps 4;

Db 106 ADYFALPRDPROSRAAEKRNVDYHNLAEIQLGAKINEFP-DQ-PTGYODGF 163
 QY 737 GAYTTRPD-RLETVIASITKPG-NITCVRRVAVQQAAGHRAKRAKSHVSTLTDLPY 794

Db 164 EKQYVAYEQNDRLHDIVASDNASYSQAMILVGM 201
 QY 795 MRQFAHLQETSPLRDVAIVIEOSSLEASSGLFDVRL 832

RESULT 10
 ENTRY S51470 #type complete
 TITLE hypothetical protein YLR383w - yeast (Saccharomyces cerevisiae)
 ALTERNATE_NAMES hypothetical protein L3502.2
 ORGANISM #formal_name Saccharomyces cerevisiae
 DATE 23-Feb-1995 #sequence_revision 12-May-1995 #text_change 06-Feb-1998
 #accessions S51470; S20170; S20168; S20163
 #cross-references S51466

REFERENCE
 #authors Du, Z.
 #submission submitted to the EMBL Data Library, December 1994
 #description The sequence of S. cerevisiae cosmid L3502.
 #accession S51470
 #molecule_type DNA
 #residues 1-1114 #label DVZ
 #cross-references EMBL:U19104; NID:9609423; PID:9609425; MIPS:YLR383w
 REFERENCE
 #journal S15581
 #authors Slonimski, P.P.
 #submission submitted to the EMBL Data Library, October 1987
 #accession S20170
 #molecule_type DNA
 #residues 1-184 #label SLO

REFERENCE
 #journal A91079
 #authors Labouesse, M.; Herbert, C.J.; Dujardin, G.; Slonimski, P.P.
 #journal EMBO J. (1987) 6:713-721
 #title Three suppressor mutations which cure a mitochondrial RNA maturase deficiency occur at the same codon in the open reading frame of the nuclear RNM2 gene.
 #cross-references MIMD:87218532
 #accession S20168
 #status translation not shown


```

##molecule_type DNA
##residues 1-61 ##label LAB
##cross-references EMBL:X05143
S18752
#authors Herbert, C.J.; DuJardin, G.; Labouesse, M.; Slonimski, P.P.
#journal Mol. Gen. Genet. (1988) 213:297-309
#title Divergence of the mitochondrial leucyl tRNA synthetase genes
in two closely related yeasts Saccharomyces cerevisiae and
Saccharomyces douglasii: a paradigm of incipient evolution.
#cross-references MIMD:89039717
#accession S20163
##status not compared with conceptual translation
##molecule_type DNA
##residues 1-62 ##label HER
GENETICS
#gene SCD:RHC18
#cross-references SCD:S0004375; MIPS:YLR383w
#map_position 12R
SUMMARY #length 1114 #molecular-weight 128008 #checksum 5560

Query Match 4.6%; Score 101; DB 2; Length 1114;
Best Local Similarity 40.8%; Pred. No. 2.04e+00;
Matches 20; Conservative 12; Mismatches 11; Indels 6; Gaps 5;

Db 45 MTCNSTLDEDSGFRSSDYATADQDNFEEPSPIYKIVLRNFMCH 93
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 795 MRFVHH-LOETSPLRDA--VVI-EQSSSLNENSSGLFD-VFLR-FMCH 837

RESULT 11
ENTRY
TITLE GNMVYP #type complete
genome polypeptide - yellow fever virus (strain Pasteur
17D-204)
CONTAINS
capsid protein C; envelope protein M; major envelope protein
E; nonstructural protein NS1; nonstructural protein NS2a;
nonstructural protein NS2b; nonstructural protein NS3;
nonstructural protein NS4a; nonstructural protein NS4b;
nonstructural protein NS5
#formal_name yellow fever virus
#ORGANISM
DATE 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change
12-Apr-1996
ACCESSIONS
S07757
S07757
REFERENCES
#authors Dupuy, A.; Despres, P.; Cahour, A.; Girard, M.; Bouloy, M.
#journal Nucleic Acids Res. (1989) 17:3989
#title Nucleotide sequence comparison of the genome of two 17D-204
yellow fever vaccines.
#cross-references MIMD:89282413
#accession S07757
#status not shown; translation not shown
##molecule_type genomic RNA
##residues 1-3411 ##label DUP
##cross-references EMBL:X15062
#note the nucleotide sequence was submitted to the EMBL Data
Library, April 1989, in computer-readable form
CLASSIFICATION
#superfamily yellow fever virus genome polypeptide
#capsid protein; envelope protein; glycoprotein; nonstructural
protein; polypeptide; transmembrane protein
KEYWORDS
FEATURE
2-210 #product capsid protein C #status predicted #label CPC\
105-125 #domain transmembrane #status predicted #label TM1\
211-285 #product envelope protein M #status predicted #label
EPN\
271-289 #domain transmembrane #status predicted #label TM2\
286-778 #product major envelope protein E #status predicted
#label ME\
736-753 #domain transmembrane #status predicted #label TM3\
756-778 #domain transmembrane #status predicted #label TM4\
779-1187 #product nonstructural protein NS1 #status predicted
#label NS1\
1133-1151 #domain transmembrane #status predicted #label TM5\
1160-1179 #domain transmembrane #status predicted #label TM6\
1188-1354 #product nonstructural protein NS2a #status predicted

```

1355-1484	#label N2A\	#product nonstructural protein NS2b	#status predicted
1485-2107	#label N2B\	#product nonstructural protein NS3	#status predicted
2108-2394	#label NS3\	#product nonstructural protein NS4a	#status predicted
2395-2506	#label N4A\	#product nonstructural protein NS4b	#status predicted
2507-3411	#label N4B\	#product nonstructural protein NS5	#status predicted
134, 150, 172, 594, 908, 986, 1176, 2062, 2320, 2346, 2408, 2467, 2720, 2734, 2740	#binding_site carbohydrate (Asn) (covalent)	#status predicted	
SUMMARY	#length 3411	#molecular_weight 379528	#checksum 8417
Query Match	4.68;	Score 101;	DB 1; Length 3411;
Best Local Similarity	32.38;	Pred. No. 2,04e+00;	
Matches	21;	Conservative	14; Mismatches 27; Indels 3; Gaps 3;
Db	1334	GLT0PFG-1CARLARIRGRSIPVNEALAGLVAGVAGIAPOEMN-FLSPINAGL	1391
QY	826	GLEPVEFRFPCDH-AVRIRKSTVQCQCGIPGSLSTLCTGDEMKNKFLACIRBDGL	884
Db	1392	LMMLV	1396
QY	885	LLRLV	889
RESULT	12		
ENTRY	GNMVT	#type complete	
TITLE	genome polyprotein C - yellow fever virus (strain 17D)		
CONTAINS	capsid protein C; envelope protein M; major envelope protein E; nonstructural protein NS1; nonstructural protein NS2a; nonstructural protein NS2b; nonstructural protein NS3; nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5		
ORGANISM	#formal_name yellow fever virus		
DATE	27-Nov-1985	#sequence_revision 27-Nov-1985	#text_change 12-Apr-1996
ACCESSIONS	A03914		
REFERENCE	A03914		
authors	Rice, C.M.; Lenches, E.M.; Eddy, S.R.; Shin, S.J.; Sheets, R.L.; Strauss, J.H.		
#journal	Science (1985) 229:726-733		
#title	Nucleotide sequence of yellow fever virus: implications for flavivirus gene expression and evolution.		
#cross-references	MOID:85272570		
#accession	A03914		
KEYWORDS	##molecule_type genomic RNA		
CLASSIFICATION	1-3411 ##label RIC		
KEYWORDS	#superfamily yellow fever virus genome polyprotein capsid protein; envelope protein; glycoprotein; nonstructural protein; polyprotein; transmembrane protein		
FEATURE			
2-210	#product capsid protein C	#status predicted	#label CPC\
211-285	#product envelope protein M	#status predicted	#label EPM\
249-269	#domain transmembrane	#status predicted	#label TM1\
271-285	#domain transmembrane	#status predicted	#label TM2\
286-278	#product major envelope protein E	#status predicted	#label MEJ\
740-753	#domain transmembrane	#status predicted	#label TM3\
755-778	#domain transmembrane	#status predicted	#label TM4\
779-1187	#product nonstructural protein NS1	#status predicted	#label NS1\
1159-1180	#domain transmembrane	#status predicted	#label TM5\
1188-1354	#product nonstructural protein NS2a	#status predicted	#label N2A\
1355-1484	#product nonstructural protein NS2b	#status predicted	

Qy	Db	Accessions	Reference	Authors	Organism	Date	Result Entry Title	Score	DB 1	Length	3411	Score	DB 2	Length	372	Score	DB 3	Length	372	Score	DB 4	Status	DB 5	Status	DB 6	Status	DB 7	Status	DB 8	Status	DB 9	Status	DB 10	Status	DB 11	Status	DB 12	Status	DB 13	Status	DB 14	Status	DB 15	Status	DB 16	Status	DB 17	Status	DB 18	Status	DB 19	Status	DB 20	Status	DB 21	Status	DB 22	Status	DB 23	Status	DB 24	Status	DB 25	Status	DB 26	Status	DB 27	Status	DB 28	Status	DB 29	Status	DB 30	Status	DB 31	Status	DB 32	Status	DB 33	Status	DB 34	Status	DB 35	Status	DB 36	Status	DB 37	Status	DB 38	Status	DB 39	Status	DB 40	Status	DB 41	Status	DB 42	Status	DB 43	Status	DB 44	Status	DB 45	Status	DB 46	Status	DB 47	Status	DB 48	Status	DB 49	Status	DB 50	Status	DB 51	Status	DB 52	Status	DB 53	Status	DB 54	Status	DB 55	Status	DB 56	Status	DB 57	Status	DB 58	Status	DB 59	Status	DB 60	Status	DB 61	Status	DB 62	Status	DB 63	Status	DB 64	Status	DB 65	Status	DB 66	Status	DB 67	Status	DB 68	Status	DB 69	Status	DB 70	Status	DB 71	Status	DB 72	Status	DB 73	Status	DB 74	Status	DB 75	Status	DB 76	Status	DB 77	Status	DB 78	Status	DB 79	Status	DB 80	Status	DB 81	Status	DB 82	Status	DB 83	Status	DB 84	Status	DB 85	Status	DB 86	Status	DB 87	Status	DB 88	Status	DB 89	Status	DB 90	Status	DB 91	Status	DB 92	Status	DB 93	Status	DB 94	Status	DB 95	Status	DB 96	Status	DB 97	Status	DB 98	Status	DB 99	Status	DB 100	Status	DB 101	Status	DB 102	Status	DB 103	Status	DB 104	Status	DB 105	Status	DB 106	Status	DB 107	Status	DB 108	Status	DB 109	Status	DB 110	Status	DB 111	Status	DB 112	Status	DB 113	Status	DB 114	Status	DB 115	Status	DB 116	Status	DB 117	Status	DB 118	Status	DB 119	Status	DB 120	Status	DB 121	Status	DB 122	Status	DB 123	Status	DB 124	Status	DB 125	Status	DB 126	Status	DB 127	Status	DB 128	Status	DB 129	Status	DB 130	Status	DB 131	Status	DB 132	Status	DB 133	Status	DB 134	Status	DB 135	Status	DB 136	Status	DB 137	Status	DB 138	Status	DB 139	Status	DB 140	Status	DB 141	Status	DB 142	Status	DB 143	Status	DB 144	Status	DB 145	Status	DB 146	Status	DB 147	Status	DB 148	Status	DB 149	Status	DB 150	Status	DB 151	Status	DB 152	Status	DB 153	Status	DB 154	Status	DB 155	Status	DB 156	Status	DB 157	Status	DB 158	Status	DB 159	Status	DB 160	Status	DB 161	Status	DB 162	Status	DB 163	Status	DB 164	Status	DB 165	Status	DB 166	Status	DB 167	Status	DB 168	Status	DB 169	Status	DB 170	Status	DB 171	Status	DB 172	Status	DB 173	Status	DB 174	Status	DB 175	Status	DB 176	Status	DB 177	Status	DB 178	Status	DB 179	Status	DB 180	Status	DB 181	Status	DB 182	Status	DB 183	Status	DB 184	Status	DB 185	Status	DB 186	Status	DB 187	Status	DB 188	Status	DB 189	Status	DB 190	Status	DB 191	Status	DB 192	Status	DB 193	Status	DB 194	Status	DB 195	Status	DB 196	Status	DB 197	Status	DB 198	Status	DB 199	Status	DB 200	Status	DB 201	Status	DB 202	Status	DB 203	Status	DB 204	Status	DB 205	Status	DB 206	Status	DB 207	Status	DB 208	Status	DB 209	Status	DB 210	Status	DB 211	Status	DB 212	Status	DB 213	Status	DB 214	Status	DB 215	Status	DB 216	Status	DB 217	Status	DB 218	Status	DB 219	Status	DB 220	Status	DB 221	Status	DB 222	Status	DB 223	Status	DB 224	Status	DB 225	Status	DB 226	Status	DB 227	Status	DB 228	Status	DB 229	Status	DB 230	Status	DB 231	Status	DB 232	Status	DB 233	Status	DB 234	Status	DB 235	Status	DB 236	Status	DB 237	Status	DB 238	Status	DB 239	Status	DB 240
----	----	------------	-----------	---------	----------	------	--------------------	-------	------	--------	------	-------	------	--------	-----	-------	------	--------	-----	-------	------	--------	------	--------	------	--------	------	--------	------	--------	------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	-------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------	--------

[illegible]

 (TM)

Release 3.1A John F. Collins, Blocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Dec 18 18:37:17 1998; Maspar time 9.66 Seconds
 Tabular output not generated. 836.466 Million cell updates/sec

Title: >US-08-951-733-14
 Description: (640-940) from US08951733.pep (2 of 2)
 Perfect Score: 2214
 Sequence: 1 TSKRLRFPKPDGLRPLVNM.....NKRKYVNPVDEALGTA 301

Scoring table:
 PAM 150
 Gap 11

Searched: 74019 seqs, 26840295 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot35
 1:swissprot

Statistics: Mean 49.063; Variance 85.433; scale 0.574

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	103	4.7	181	1	YMD6_YEAST	2.00e-01
2	105	4.7	890	1	ACOL1_ECOLI	1.06e-01
3	103	4.7	891	1	ACON1_ECOLI	2.00e-01
4	102	4.6	205	1	YAO1_BPH1	2.75e-01
5	101	4.6	551	1	MCPL1_ECOLI	3.76e-01
6	101	4.6	1114	1	RH18_YEAST	3.76e-01
7	101	4.6	3411	1	POLG_YEAV1	3.76e-01
8	101	4.6	3411	1	POLG_YEAV2	3.76e-01
9	99	4.5	446	1	MOXY_PARDE	7.00e-01
10	100	4.5	1030	1	YEO2_SCHRO	5.14e-01
11	98	4.4	586	1	YCE1_YEAST	9.52e-01
12	98	4.4	906	1	ACON1_ECOLI	9.52e-01
13	97	4.4	1238	1	PTPJ_MOUSE	1.29e+00
14	96	4.3	318	1	VP23_HSV1	1.74e+00
15	95	4.3	362	1	ALFC_CARAU	2.35e+00
16	96	4.3	417	1	ASPG_METJA	1.74e+00
17	96	4.3	789	1	Y1010_MYCN	5.67e+00
18	92	4.2	212	1	Y1010_MYCN	4.24e+00
19	93	4.2	261	1	Y1010_MYCN	4.24e+00
20	92	4.2	455	1	Y1010_MYCN	5.67e+00
21	92	4.2	647	1	Y1010_MYCN	5.67e+00
22	94	4.2	781	1	Y1010_MYCN	5.67e+00
23	94	4.2	901	1	Y1010_MYCN	5.67e+00

24	92	4.2	1117	1	YDVG_SCHRO	PUTATIVE HELICASE C6F1	5.67e+00
25	91	4.1	124	1	RR12_EUGR	CHLOROPLAST 30S RIBOSO	7.56e+00
26	91	4.1	164	1	PHEA_PREDI	C-PHYCOERYTHRIN ALPHA	7.56e+00
27	90	4.1	297	1	YDVG_BACSD	HYPOTHETICAL 32.5 KD P	1.00e+01
28	90	4.1	309	1	VIRB_SHIFT	VIRULENCE REGULON TRAN	1.00e+01
29	90	4.1	357	1	ALF_SPIOL	FRUCTOSE-BISPHOSPHATE	1.00e+01
30	91	4.1	366	1	YDVG_ECOLI	HYPOTHETICAL 40.4 KD P	7.56e+00
31	90	4.1	420	1	MLIC_XENLA	MELATONIN RECEPTOR TYP	1.00e+01
32	90	4.1	428	1	DD11_YEAST	DNA-DAMAGE INDUCIBLE P	1.00e+01
33	90	4.1	433	1	FUSAR_BURCE	FUSARIC ACID RESISTANC	1.00e+01
34	90	4.1	445	1	LAEB_VIBPA	FLAGELLAR HOOK-ASSOCIA	1.00e+01
35	90	4.1	452	1	YHA2_YEAST	HYPOTHETICAL 51.2 KD P	1.00e+01
36	91	4.1	609	1	YVRC_YEAST	EXCINCLASE ABC SUBUN	7.56e+00
37	91	4.1	695	1	PIP6_BOVIN	1-PHOSPHATIDYLINOSITOL	7.56e+00
38	91	4.1	780	1	K6PL_MOUSE	6-PHOSPHORUCTOKINASE,	1.00e+01
39	90	4.1	1475	1	TRA2_CAEEL	SEX-DETERMINING TRANS	7.56e+00
40	91	4.1	1590	1	GCN2_YEAST	PROTEIN KINASE GCN2 (E	7.56e+00
41	89	4.0	124	1	RR12_HORVU	CHLOROPLAST 30S RIBOSO	1.33e+01
42	89	4.0	127	1	TRM8_ECOLI	TRAM PROTEIN.	1.33e+01
43	89	4.0	133	1	RR12_CHLRE	CHLOROPLAST 30S RIBOSO	1.33e+01
44	89	4.0	532	1	RTCR_ECOLI	TRANSCRIPTIONAL REGULA	1.33e+01
45	89	4.0	957	1	NIRB_KLEBN	NITRITE REDUCTASE (NAD	1.33e+01

ALIGNMENTS

RESULT	ID	YMD6_YEAST	STANDARD	PRT	181 AA.
AC	003705				
DT	01-NOV-1997 (REL. 35, CREATED)				
DT	01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)				
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)				
DE	HYPOTHETICAL 20.7 KD PROTEIN IN CAT2-AND1 INTERGENIC REGION.				
GN	YMD036W.				
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).				
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRATIN-S288C / AB972;				
RA	BADCOCK K., CHURCHER C., BARRELL B.G., RAJANDREAM M.A., WALSH S.V.;				
RL	SUBMITTED (NOV-1994) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR	EMBL: Z46593; G575683; ALT_SEQ.				
KW	HYPOTHETICAL PROTEIN.				
QY	SEQUENCE 181 AA; 20662 MW; AC111C34 CRC32;				
DB	50 GEOMYSALYKTLIEV-KYMKRTNLSSECVLCSPNSISDAE-LKFGIKDSSOLICL 107				
QY	673 AERLTISRV-KALFESVLYNERARRPGLLGASVLDGDIHRAMRTFLVLRADPPPELYFV 731				
DB	108 KFH-TNT-DVYDKRQLRTIMTSYKGOE 133				
QY	732 KVDVTGAYDITIPDRLEVIASIIKPN 759				
RESULT	2				
ID	ACOL1_ECOLI	STANDARD	PRT	890 AA.	
AC	P25516; P78060; P78148;				
DT	01-MAY-1992 (REL. 22, CREATED)				
DT	01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)				
DT	01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)				
DE	ACONITATE HYDRAEASE 1 (EC 4.2.1.3) (CITRATE HYDRO-LYASE 1)				
DE	(ACONITATE 1).				
GN	ACNA OR ACN.				
OS	ESCHERICHIA COLI.				
OC	PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;				
OC	ENTEROBACTERIACEAE.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE; 92174916.				

RA PRODRMOMU C., ARTYMIUK P.J., GUEST J.R.;
 RL EUR. J. BIOCHEM. 204:599-609(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RA BLATTNER F.R., PLUNKETT G. III, MAYHEW G.F., PERNA N.T., GLASNER F.D.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA AIBA H., BABA T., FUJITA K., HAYASHI K., HONJO A., HORIUCHI T.,
 RA IKEMOTO K., INADA T., ISONO K., ISONO S., ITOH T., KANAI K.,
 RA KASAI H., KASHIMOTO K., KIM S., KIMURA S., KITAGAWA M.,
 RA KITAHARA M., MAKINO K., MASUDA S., MIKI T., MIZOBUCHI K., MORI H.,
 RA MOTOKURA K., NAKAMURA Y., NASHIMOTO H., NISHIO Y., OSHIMA T.,
 RA SAITO N., SAMPEI G., SEKI Y., TAGAMI H., TAKEMOTO K., WADA C.,
 RA YAMAMOTO Y., YANO M.;
 RL SUBMITTED (DEC-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [4]
 RP SEQUENCE OF 1-18.
 RX MEDLINE: 92148368.
 RA PRODRMOMU C., HAYNES M.J., GUEST J.R.;
 RL J. GEN. MICROBIOL. 137:2505-2515(1991).
 CC -1- FUNCTION: MAY HAVE AN IRON-RESPONSIVE REGULATORY FUNCTION.
 CC -1- CATALYTIC ACTIVITY: CITRATE - CIS-ACONITATE + H(2)O.
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
 CC -1- SUBUNIT: MONOMER.
 CC -1- ACONITATE HAS AN ACTIVE (4FE-4S) AND AN INACTIVE (3FE-4S) FORMS.
 CC THE ACTIVE (4FE-4S) CLUSTER IS PART OF THE CATALYTIC SITE THAT
 CC INTERCONVERTS CITRATE, CIS-ACONITATE, AND ISOCITRATE.
 CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
 DR EMBL: A6000225; G187531; -;
 DR EMBL: D90765; G1742087; -;
 DR EMBL: D90766; G1742090; -;
 DR PIR: S22375; S22375.
 DR ECOGENE: EG11325; ACNA.
 DR PROSITE: PS00450; ACONITASE_1; 1.
 DR PROSITE: PS01244; ACONITASE_2; 1.
 KW LYASE: TRICARBOXYLIC ACID CYCLE; IRON-SULFUR; 4FE-4S.
 FT INIT_MET 0
 FT METAL 434 434
 FT METAL 500 500
 FT METAL 503 503
 FT METAL 503 503
 FT METAL 521 521
 FT CONFLICT 521 521 G -> S (IN REF. 2).
 SQ SEQUENCE 890 AA; 97516 MW; 276673F6 CRC32;
 Query Match 4.7%; Score 105; DB 1; Length 890;
 Best local Similarity 21.9%; Pred. No. 1,06e-01;
 Matches 14; Conservative 22; Mismatches 25; Indels 3; Gaps 3;
 Db 754 MRYKQEQTPLAVIAGKYGSSSR-DMAAGPRLG-IRVIVAE-SFERIHSNLTGMGI 810
 Qy 643 LRIPKPDGLPIVNDYVVGARTFRERKARLTSTVKALFSLNTERARRPGLGASV 702
 Db 811 LPLE 814
 Qy 703 LGLD 706
 RESULT 3
 ID ACON_LEGPN STANDARD; PRT; 891 AA.
 AC P37032;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
 DE ACONITASE HYDRAITASE (EC 4.2.1.3) (CITRATE HYDRO-LYASE) (ACONITASE)
 DE (MAJOR IRON-CONTAINING PROTEIN) (MCP) (IP210).
 GN ACN.
 OS LEGIONELLA PNEUMOPHILA.

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 OC LEGIONELLACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-PHILADELPHIA 1;
 RX MEDLINE: 93374864.
 RA MENGAUD J.M., HORWITZ M.A.;
 RL BACTERIOL. 175:5666-5676(1993).
 CC -1- CATALYTIC ACTIVITY: CITRATE - CIS-ACONITATE + H(2)O.
 CC -1- PATHWAY: TRICARBOXYLIC ACID CYCLE.
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
 CC -1- ACONITASE HAS AN ACTIVE (4FE-4S) AND AN INACTIVE (3FE-4S) FORMS.
 CC THE ACTIVE (4FE-4S) CLUSTER IS PART OF THE CATALYTIC SITE THAT
 CC INTERCONVERTS CITRATE, CIS-ACONITATE, AND ISOCITRATE.
 CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY.
 DR EMBL: L22081; G348945; -;
 DR PIR: B48642; B48642.
 DR HSSP: P20004; IACO.
 DR PROSITE: PS00450; ACONITASE_1; 1.
 DR PROSITE: PS01244; ACONITASE_2; 1.
 KW LYASE: TRICARBOXYLIC ACID CYCLE; IRON-SULFUR; 4FE-4S.
 FT METAL 435 435
 FT METAL 501 501
 FT METAL 501 501
 FT METAL 504 504
 FT METAL 504 504
 SQ SEQUENCE 891 AA; 98209 MW; E6230AC5 CRC32;
 Query Match 4.7%; Score 103; DB 1; Length 891;
 Best local Similarity 25.0%; Pred. No. 2,00e-01;
 Matches 16; Conservative 22; Mismatches 23; Indels 3; Gaps 3;
 Db 755 MRYQEQDPLVIAKRYGTSSR-DMAAGTNLLG-VKAVITE-SFERIHSNLTGMGI 811
 Qy 643 LRIPKPDGLPIVNDYVVGARTFRERKARLTSTVKALFSLNTERARRPGLGASV 702
 Db 812 LPLQ 815
 Qy 703 LGLD 706
 RESULT 4
 ID Y001_BPHPI STANDARD; PRT; 205 AA.
 AC P51700;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 23.2 KD PROTEIN IN INT-CL INTERGENIC REGION (ORF1).
 DE BACTERIOPHAGE HPI.
 OS VIRIDAE; DS-DNA NONENVELOPED VIRUSES; MYOVIRIDAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HP1C1;
 RX MEDLINE: 95089704.
 RA ESPOSITO D., SCOCIA J.J.;
 RL MOL. MICROBIOL. 13:685-695(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HP1C1;
 RX MEDLINE: 96279738.
 RA ESPOSITO D., FITZMAURICE W.P., BENJAMIN R.C., GOODMAN S.D.,
 RA WALDMAN A.S., SCOCIA J.J.;
 RL NUCLEIC ACIDS RES. 24:2360-2368(1996).
 DR EMBL: U24159; G459176; -;
 KW HYPOTHETICAL PROTEIN; TRANSMEMBRANE.
 FT TRANSMEM 12 69
 FT TRANSMEM 49 32
 FT SEQUENCE 205 AA; 23260 MW; 821C7A9F CRC32;
 Query Match 4.6%; Score 102; DB 1; Length 205;
 Best local Similarity 22.1%; Pred. No. 2,75e-01;
 Matches 33; Conservative 49; Mismatches 50; Indels 17; Gaps 17;

[illegible]

Query Match	4.68;	Score 101;	DB 1;	Length 3411;
Best Local Similarity	32.38;	Pred. No. 3.76e-01;		
Matches	21;	Conservative	14;	Mismatches 27;
			Indels	3;
			Gaps	3;

RESULT	9	
ID	MOXY_PARDE	STANDARD; PRT; 446 AA

OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI
OC UNCERTAIN.
OC

CC -1- FUNCTION: MEMBER OF THE TWO-COMPONENT REGULATORY SYSTEM MOXY/MOXX
CC PROBABLY INVOLVED IN THE REGULATION OF THE METHANOL DEHYDROGENASE
CC EXPRESSION. MAY FUNCTION AS A MEMBRANE-ASSOCIATED PROTEIN KINASE
CC THAT PHOSPHORYLATES MOXY IN RESPONSE TO ENVIRONMENTAL SIGNALS.

CC -1 SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE (PROBABLE).
CC -1 SIMILARITY: TO OTHER PROKARYOTIC SENSORY TRANSDUCTION HISTIDINE KINASES.

DR EMBL; M92421; G150599; -.
DR PIR; S32871; S32871.
KW SENSOR TRANSDUCTION; TRANSFERASE; KINASE; PHOSPHORYLATION
KW TRANSMEMBRANE; INNER MEMBRANE; METANOL UTILIZATION.

FT	TRANSMEM	1	21	POTENTIAL.
FT	TRANSMEM	101	121	POTENTIAL.
FT	TRANSMEM	144	164	POTENTIAL.
SQ	SEQUENCE	446 AA;	48236 MM;	D59FB521 CRC32

[illegible]

RESULT	10	
ID	YE02_SCHPO	STANDARD;
AC	013788.	PRT; 1030 AA.

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA SHELTON J., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.,

RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DDAJ DATA BANKS.
CC -|- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).
CC -|- SIMILARITY: BELONGS TO THE SKI2 SUBFAMILY OF HELICASES.
DR EMBL, Z98597; E334274; -.
KW HYPOTHETICAL PROTEIN; HELICASE; ATP-BINDING; NUCLEAR PROTEIN.
FT NP_BIND 147 154 ATP (POTENTIAL).
FT SITE 238 241 DEHY BOX.
SQ SEQUENCE 1030 AA; 118460 MW; C334C59C CRC32;

Db 116 DPpAKTyPEfELDPFOStAlNCVERMESVLSAhtSAGkTVIAEYAlAQ-ALNRRQVlyT 174
 |||: | :| : ||: | :| : ||: | :| : | :| :
 0Y 723 DPpELyFvKVDyTGADdI.PDORLLEvYASlIKPONTyCVRRyAVOQAAGhYKKAfK 782

Db 175 SPIKSLN-QKY-RELISFEGDVLMTGDSVINSPAS 209
| : : : | : : : : : | : : |
Oy 783 SHVSTLDQPYMQFVAHQERSPLRDAYVIRQSSS 819

RESULT	11
ID	YCE1_YEAST
	STANDARD;
	PRT;
	586 AA

AC P25579; 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE HYPOTHETICAL 65.1 KD PROTEIN IN PBN1-AVAL INTERGENIC REGION.
 GN YC51W.
 OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
 CC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCETES.

RN [1]
RP SEQUENCE FROM N.A.
RA FULLER L.J., KELLY A., LEWIS C., MCKEE R.A., PEARSON B.M.;
RL SUBMITTED (MAR-1992) TO EMBL/GENBANK/DBJ DATA BANKS.

DR EMBL; X59720; E264527; -.
DR PIR: S19381; S19381.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 586 AA; 65152 MW; 56434E7A CRC32;

Query Match 4.48; Score 98; DB 1; Length 586

744 QDR-LTEVIASIIKPQNTYCVRRAYVQKAHGHVR-KAFKSHVSTLTDLQ 793

```

RESULT 14
ID VP23_HSV11 STANDARD; PRT; 318 AA.
AC P10202;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-DEC-1992 (REL. 24, LAST ANNOTATION UPDATE)
DE CAPSID PROTEIN VP23.
OS
OS HERPES SIMPLEX VIRUS (TYPE 1 / STRAIN 17).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE; 88274327.
RA MCEVOCH D.J., DUKRIMPLE M.A., DAVIDSON A.J., DOLAN A., FRAME M.C.,
RA MCNAB D., PERRY L.J., SCOTT J.E., TAYLOR P.;
RL J. GEN. VIROL. 69:1531-1574(1988).
RN [2]
RN FUNCTION.
RP MEDLINE; 90264854.
RA RIXON F.J., DAVIDSON M.D., DAVIDSON A.J.;
RL J. GEN. VIROL. 71:1211-1214(1990).
RN [3]
RN SEQUENCE OF 1-13.
RP MEDLINE; 93019027.
RA DAVIDSON M.D., RIXON F.J., DAVIDSON A.J.;
RL J. GEN. VIROL. 73:2709-2713(1992).
CC -1- FUNCTION: MAY BE A COMPONENT OF TRIMERIC STRUCTURES LINKING
CC ADJACENT CAPSOMERES AT THE CAPSID SURFACE.
CC -1- SIMILARITY: TO OTHER HERPESVIRUSES CAPSID PROTEIN VP23.
DR EMBL; D10879; G221739; -
DR EMBL; X14112; G59518; -
DR EMBL; X14112; E312332; -
DR PIR; I30083; WMBE78.
KW COAT PROTEIN
SQ SEQUENCE 318 AA; 34270 MW; FFEBO01F CRC32;

```

[illegible]

Search completed: Fri Dec 18 18:37:42 1998
Job time : 25 secs.

RESULT	15	STANDARD;	PRT;	362 AA.
ID	ALFC_CARAV			
AC	P53448:			
DT	01-OCT-1996 (REL. 34, CREATED)			
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	FRUCTOSE-BISPHOSPHATE ALDOLASE C (EC 4.1.2.13) (BRAIN).			
GN	ALDOC.			
OS	CARASSIUS AURATUS (GOLDFISH).			
OC	ENHAROTLA; METAOLA; CHORDATA; VERTEBRATA; PISCES; GNATHOSTOMATA;			
OC	OSTEICHTHYES; ACTINOPTERYGII; CYPRINIFORMES.			
FN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-BRAIN:			
RA	REICHERT T.2.; WILLIAMS M.D.; CALLARD G.V.; TOLAN D.R.;			
RL	SUBMITTED (JAN-1996), TO EMBL/GENBANK/DBJ DATA BANKS.			
CC	-1- CATALYTIC ACTIVITY: D-FRUCTOSE 1,6-BISPHOSPHATE - GLYCERONE-			
CC	PHOSPHATE + D-GICERALDEHIDE + 3-PHOSPHATE.			
CC	-1- PATHWAY: SIXTH STEP IN GLYCOLYSIS.			
CC	-1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).			

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

WIDEORIT (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_p protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:39:23 1998; MasPar time 3.38 Seconds
Tabular output not generated. 629.739 Million cell updates/sec

Title: >US-08-951-733-14
Description: (640-940) from US08951733.pep (2 of 2)
Perfect Score: 2214
Sequence: 1 TSTRIRFPKPDGLRPVNM.....NLKRTVNFVEDEALGCTA 301
Scoring table: PAM 150
Gap 11

Searched: 77309 seqs, 7078906 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5_COMB 2:PCT9_COMB 3:backfiles1

Statistics: Mean 31.637; Variance 143.394; scale 0.221

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	97	4.4	903	1	US-08-750-Sequence 1, Applicatio	6.92e+00
2	97	4.4	1398	1	US-08-750-Sequence 9, Applicatio	6.92e+00
3	96	4.3	785	1	US-08-526-Sequence 4, Applicatio	8.14e+00
4	90	4.1	433	1	US-07-661-Sequence 2, Applicatio	2.13e+01
5	85	3.8	880	2	PCT-US94-0-Sequence 5, Applicatio	4.65e+01
6	85	3.8	890	1	US-08-145-Sequence 5, Applicatio	4.65e+01
7	82	3.7	95	1	US-08-202-Sequence 23, Applicatio	7.38e+01
8	81	3.7	858	2	PCT-US93-0-Sequence 2, Applicatio	8.59e+01
9	81	3.7	1178	1	US-08-199-Sequence 2, Applicatio	8.59e+01
10	81	3.7	1178	2	PCT-US95-0-Sequence 2, Applicatio	8.59e+01
11	81	3.7	1713	2	PCT-US94-1-Sequence 24, Applicatio	8.59e+01
12	79	3.6	567	1	US-08-457-Sequence 4, Applicatio	1.16e+02
13	79	3.6	567	1	US-08-458-Sequence 4, Applicatio	1.16e+02
14	79	3.6	567	1	US-08-258-Sequence 4, Applicatio	1.16e+02
15	79	3.6	567	1	US-08-457-Sequence 4, Applicatio	1.16e+02
16	79	3.6	567	1	US-08-457-Sequence 4, Applicatio	1.16e+02
17	79	3.6	567	1	US-08-456-Sequence 4, Applicatio	1.16e+02
18	79	3.6	604	1	US-08-328-Sequence 12, Applicatio	1.16e+02
19	80	3.6	791	1	US-08-394-Sequence 2, Applicatio	1.00e+02
20	80	3.6	890	1	US-08-472-Sequence 8, Applicatio	1.00e+02
21	79	3.6	921	1	US-08-818-Sequence 2, Applicatio	1.16e+02
22	79	3.6	921	1	US-08-396-Sequence 2, Applicatio	1.16e+02
23	80	3.6	1247	1	US-08-472-Sequence 10, Applicatio	1.00e+02

RESULT	ID	US-08-750-532-1	STANDARD:	PRT:	903 AA.
24	80	3.6	3218	1	US-08-764-Sequence 27, Applicati
25	77	3.5	136	3	5212286-4 Patent No. 5212286.
26	77	3.5	146	1	US-08-453-Sequence 38, Applicati
27	77	3.5	146	1	US-08-268-Sequence 38, Applicati
28	77	3.5	151	3	5212286-2 Patent No. 5212286.
29	78	3.5	229	2	PCT-US96-0-Sequence 62, Applicati
30	78	3.5	229	2	PCT-US96-0-Sequence 13, Applicati
31	77	3.5	296	2	PCT-US95-0-Sequence 2, Applicatio
32	77	3.5	405	1	US-07-994-Sequence 2, Applicatio
33	77	3.5	405	1	US-08-421-Sequence 2, Applicatio
34	77	3.5	405	1	US-07-829-Sequence 2, Applicatio
35	77	3.5	419	1	US-08-553-Sequence 3, Applicatio
36	77	3.5	440	1	US-08-459-Sequence 2, Applicatio
37	77	3.5	619	1	US-07-762-Sequence 2, Applicatio
38	77	3.5	644	2	PCT-US93-0-Sequence 21, Applicati
39	77	3.5	647	2	PCT-US93-0-Sequence 11, Applicati
40	77	3.5	647	2	PCT-US93-0-Sequence 23, Applicati
41	77	3.5	662	1	US-08-261-Sequence 7, Applicatio
42	78	3.5	699	2	PCT-US94-0-Sequence 39, Applicati
43	77	3.5	735	2	PCT-US93-0-Sequence 13, Applicati
44	78	3.5	1019	1	US-08-271-Sequence 7, Applicatio
45	78	3.5	1612	2	PCT-US94-0-Sequence 48, Applicati

ALIGNMENTS

Sequence 1, Application US/08750532
Patent No. 5756339
GENERAL INFORMATION:
APPLICANT: MITTA, Masanori
APPLICANT: YAMAMOTO, Katsuhiko
APPLICANT: MORISHITA, Mio
APPLICANT: ASADA, Kiyozo
APPLICANT: TSUNASAWA, Susumu
APPLICANT: KATO, Ikunoshita
TITLE OF INVENTION: HYPERHERMOSTABLE PROTEASE GENE
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESSES:
ADDRESSEE: BROWDY AND NEIMARK, P. L.L.C.
STREET: 419 Seventh Street N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: United States of America
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/750,532
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP95/01095
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1994/130236
FILING DATE: 13-JUN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1994/173912
FILING DATE: 26-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.

	TELEPHONE:	(202) 628-5197	
CC	TELEFAX:	(202) 737-3528	
CC	INFORMATION FOR SEQ ID NO:	9:	
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH:	1398 amino acids	
CC	TYPE:	amino acid	
CC	STRANDEDNESS:	single	
CC	TOPOLOGY:	linear	
CC	MOLECULE TYPE:	peptide	
SQ	SEQUENCE	1398 AA; 154545 MW; 10958156 CN;	
Dd	Query Match	4.4%; Score 97; DB 1; Length 1398;	
Oy	Best Local Similarity	29.3%; Pred. No. 6,92e+00;	
Matches	17; Conservative	15; Mismatches 22; Indels 4; Gaps 4.	
Bd	743 FVLARKYDEGLEFLGYVGRRIIDPTTPVIIDEILNTIYPEKFTDENNTLTWYDI	800	
	: : : : : :	:	
Oy	745 FLVRRAQG-DP-PPELVFKVDVTGAYDGT-I PQRLT-EVIIASIIKKPNNTCYVARRAY	768	
RESULT	3	STANDARD:	PRT; 785 AA.
ID	US-08-526-964-4		
AC	xxxxxx		
XX			
DX			
DT			
Sequence	4, Application US/08526964		
DE			
CC	Sequence 4, Application US/08526964		
CC	Patent No. 5698421		
CC	GENERAL INFORMATION:		
CC	APPLICANT: Lambowitz, Alan M		
CC	APPLICANT: Zimmerly, Steven		
CC	APPICANT: Guo, Huatao		
CC	APPLICANT: Yang, Jian		
CC	TITLE OF INVENTION: Nucleotide Integrase Preparation		
CC	NUMBER OF SEQUENCES: 17		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Calfee, Halter & Griswold		
CC	STREET: 800 Superior Avenue		
CC	CITY: Cleveland		
CC	STATE: Ohio		
CC	COUNTRY: USA		
CC	ZIP: 44114		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: floppy disk		
CC	COMPUETER: IBM PC compatible		
CC	OPERATING SYSTEM: PC/DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/526,964		
CC	FILING DATE:		
CC	CLASSIFICATION: 435		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: Golrick, Mary E		
CC	REGISTRATION NUMBER: 34,829		
CC	REFERENCE/DOCKET NUMBER: 22727/00127		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (216) 622-8200		
CC	TELEFAX: (216) 241-0816		
CC	INFORMATION FOR SEQ ID NO: 4:		
CC	SEQUENCE CHARACTERISICS:		
CC	LENGTH: 785 amino acids		
CC	TYPE: amino acid		
CC	STRANDEDNESS: single		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: peptide		
CC	HYPOTHETICAL: NO		
SQ	SEQUENCE	785 AA; 90390 MW; 3218308 CN;	
Query Match	4.3%;	Score 96; DB 1; Length 785;	
Best Local Similarity	44.4%;	Pred. No. 8.14e+00;	

[illegible]

```

CC      CITY: Lexington
CC      STATE: MA
CC      COUNTRY: USA
CC      ZIP: 02173
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentin Release #1.05, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/202,389
CC      FILING DATE: 28-FEB-1994
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/983,926
CC      FILING DATE: 01-DEC-1992
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/829,141
CC      FILING DATE: 31-JAN-1992
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/721,112
CC      FILING DATE: 26-JUN-1991
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Granahan, Patricia
CC      REGISTRATION NUMBER: 32,227
CC      REFERENCE/DOCKET NUMBER: BIH92-05WA
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (617) 861-6240
CC      TELEFAX: (617) 861-9540
CC      INFORMATION FOR SEQ ID NO: 23:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 95 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 95 AA; 1151 MW; 48023 CN;
SQ
Query Match 3.7%; Score 82; DB 1; Length 95;
Best Local Similarity 22.9%; Pred. No. 7.38e+01;
Matches 16; Conservative 23; Mismatches 28; Indels 3; Gaps 3;
Db 16 YV-VDPPIAEVN-MPOYKREFEKVDARDGSSRTVQFOFIDWPEOG-VPKSGEGFIDFI 72
QY 729 YFVKVDVIGADITIQDRLTEVIASIIKPNQYICVRRAYVQKAAIGHAKAKSHVSTL 788
Db 73 GQVHKTKEQF 82
QY 789 TDLQPYMRQF 798
RESULT 8
ID PCT-US93-07213-2 STANDARD; PRT; 858 AA.
XX
XX xxxxxx
XX
DE Sequence 2, Application PC/TUS9307213
XX
CC Sequence 2, Application PC/TUS9307213
CC GENERAL INFORMATION:
CC APPLICANT: The Government of the United States of
CC APPLICANT: America, as represented by The Secretary
CC TITLE OF INVENTION: NUCLEIC ACIDS AND AMINO ACID
CC TITLE OF INVENTION: SEQUENCES FOR A STRONGLY IMMUNOREACTIVE PROTEIN ENCODED
CC TITLE OF INVENTION: HUMAN HERPESVIRUS 6 STRAIN Z29
CC NUMBER OF SEQUENCES: 16
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: The Government of the United States of
CC ADDRESSEE: America, as represented by The Secretary
CC STREET: 6011 Executive Blvd., Suite 325
CC CITY: Rockville
CC STATE: Maryland
CC

```


CC APPLICANT: Beck, James Joseph
CC APPLICANT: Hill, Dwight Steven
CC APPLICANT: Ryals, John Andrew
CC APPLICANT: Gaffney, Thomas Deane
CC APPLICANT: Lam, Stephen Ting
CC APPLICANT: Hammer, Phillip E.
CC APPLICANT: Uknes, Scott Joseph
CC TITLE OF INVENTION: Genes for the synthesis of
CC TITLE OF INVENTION: antipathogenic substances
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ciba-Geigy Corporation
CC STREET: 7 Skyline Drive
CC CITY: Hawthorne
CC STATE: NY
CC COUNTRY: USA
CC ZIP: 10532
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/458,076A
CC FILING DATE: 01-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/457,205
CC FILING DATE: 01-JUN-1995
CC APPLICATION NUMBER: 08/258,261
CC FILING DATE: 08-JUN-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elmer, James Scott
CC REGISTRATION NUMBER: 36,129
CC REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 919-541-8614
CC TELEFAX: 919-541-8689
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 567 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 567 AA; 65037 MW; 1552397 CN;
SQ
Query Match 3.6%; Score 79; DB 1; Length 567;
Best Local Similarity 25.0%; Pred. No. 1.16e+02;
Matches 8; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
Db 141 LLOAIIYKGVKHOHTVTEYHADKGVAVTT 172
Y 909 LVKGVPEYGVNLRKTVNFPVEDALGCTA 940
RESULT 14
ID US-08-258-261B-4 STANDARD; PRT; 567 AA.
XX xxxxxx
DE Sequence 4, Application US/08258261B
CC Patent No. 5639949
CC GENERAL INFORMATION:
CC APPLICANT: Schnupp, Thomas
CC APPLICANT: Ligon, James M.
CC APPLICANT: Beck, James Joseph
CC APPLICANT: Hill, Dwight Steven
CC APPLICANT: Ryals, John Andrew
CC APPLICANT: Gaffney, Thomas Deane

CC APPLICANT: Lam, Stephen Ting
CC APPLICANT: Hammer, Phillip E.
CC APPLICANT: Uknes, Scott Joseph
CC TITLE OF INVENTION: Genes for the synthesis of
CC TITLE OF INVENTION: antipathogenic substances
CC NUMBER OF SEQUENCES: 22
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ciba-Geigy Corporation
CC STREET: 7 Skyline Drive
CC CITY: Hawthorne
CC STATE: NY
CC COUNTRY: USA
CC ZIP: 10532
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/258,261B
CC FILING DATE: 08-JUN-1994
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/457,205
CC FILING DATE: 01-JUN-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elmer, James Scott
CC REGISTRATION NUMBER: 36,129
CC REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 919-541-8614
CC TELEFAX: 919-541-8689
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 567 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 567 AA; 65037 MW; 1552397 CN;
SQ
Query Match 3.6%; Score 79; DB 1; Length 567;
Best Local Similarity 25.0%; Pred. No. 1.16e+02;
Matches 8; Conservative 13; Mismatches 11; Indels 0; Gaps 0;
Db 141 LLOAIIYKGVKHOHTVTEYHADKGVAVTT 172
Y 909 LVKGVPEYGVNLRKTVNFPVEDALGCTA 940
RESULT 15
ID US-08-457-342-4 STANDARD; PRT; 567 AA.
XX xxxxxx
DE Sequence 4, Application US/08457342
CC Patent No. 5652898
CC GENERAL INFORMATION:
CC APPLICANT: Schnupp, Thomas
CC APPLICANT: Ligon, James M.
CC APPLICANT: Beck, James Joseph
CC APPLICANT: Hill, Dwight Steven
CC APPLICANT: Ryals, John Andrew
CC APPLICANT: Gaffney, Thomas Deane
CC APPLICANT: Lam, Stephen Ting
CC APPLICANT: Hammer, Phillip E.
CC APPLICANT: Uknes, Scott Joseph
CC TITLE OF INVENTION: Genes for the synthesis of
CC TITLE OF INVENTION: antipathogenic substances
CC NUMBER OF SEQUENCES: 22

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Ciba-Geigy Corporation
CC STREET: 7 Skyline Drive
CC CITY: Hawthorne
CC STATE: NY
CC COUNTRY: USA
CC ZIP: 10532
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,342
CC FILING DATE: 01-JUN-1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/457,205
CC FILING DATE: 01-JUN-1995
CC APPLICATION NUMBER: 08/258,261
CC FILING DATE: 08-Jun-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Elmer, James Scott
CC REGISTRATION NUMBER: 36,129
CC REFERENCE/DOCKET NUMBER: CGC 1506/CIP3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 919-541-8614
CC TELEFAX: 919-541-8689
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 567 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: Protein
CC SEQUENCE 567 AA; 65037 MW; 1552397 CN;

Query Match 3.6%; Score 79; DB 1; Length 567;
Best Local Similarity 25.0%; Pred. No. 1.16e+02;
Matches 8; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

Db 141 LLOAIIKYGCKVHQKTVEYHADRKGAVAT 172
|::: |||::: ||::: |:
QY 909 LVKGVPEYGCNVNLRKTVNFPVEDALGGTA 940

Search completed: Fri Dec 18 18:39:40 1998
Job time : 17 secs.

QY 239 PAPGARRRGGSASRS-LPLPKRR-RCGAPEPERTVGGSGMAHGRTRGSPSD-RGFCV 295
 Db 96 tppappedgrrggaagaastrgrgrrpprppskappxerw 140
 QY 296 SPAPPAEATSLGALSGTRHSHPSVGRQHHAGPSTSRP-PRW 339

RESULT 2

ID W18664 standard; Protein; 386 AA.

AC W18664; 24-JUL-1997 (first entry)
 DE Fragmented human NF-H gene +1 frameshift mutant product.
 KW Frameshift mutation product; GAGA motif; somatic mutation; diagnosis;
 KW detection; antibody; probe; cancer; neoplasia; neurodegenerative;
 KW Parkinson's; Alzheimer's disease; Pick's; Huntington's disease;
 KW Down's syndrome; frontal lobe dementia; progressive supranuclear palsy;
 KW PSP; amyotrophic lateral sclerosis; multiple sclerosis; MS;
 KW cardiovascular; rheumatoid arthritis.
 OS Homo sapiens.

FT Key Location/Qualifiers

FT misc_difference 1..387 /note="X corresponds to a stop codon in the
 FT accompanying DNA file, T69796"

FT peptide /note="antigenic peptide used for antibody
 FT production"

FT MO9712993-A2.

PN 10-APR-1997.

PD 02-OCT-1996; IB1106.

PF 02-OCT-1995; GB-020080.

PR 11-JAN-1996; US-009832.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

PA (UYUR-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

DR WPI; 97-226235/20.

DR N-PSDB; T69796.

PT Use of mutant genes having frame:shift mutation(s) - for developing
 PT prods. for the diagnosis, prevention and treatment of associated

PT diseases, e.g. cancer or neurodegenerative disease

PS Claim 22; Fig 9; 13pp; English.

CC W18663 and W18664 are +2 and +1 frameshift mutations, respectively,
 CC of a sequence comprising fragments of the coding sequence of the

CC human neurofilament subunit NF-H gene corresponding to nucleotides
 CC 1-1162 of the wild-type NF-H gene. This region contains GAGAG motifs.

CC Frameshift mutants of the tau, ubiquitin, apolipoprotein E,
 CC microtubule-associated protein 2 (MAP-2), neurofilament subunit I, M

CC and H and amyloid A4 genes are claimed. All these genes share a common
 CC GAGAN motif (N-A, G, C or T), which is the site of common GA

CC dinucleotide deletion(s) that cause neurodegenerative disorders.

CC Antigenic peptides used for the production of antibodies, and small
 CC nucleic acid sequences derived from frameshift mutants are used in the

CC diagnosis, prevention and treatment of cancer and neurodegenerative
 CC diseases, e.g. Parkinson's disease, Alzheimer's disease, Down's

CC syndrome, frontal lobe dementia (Pick's disease), progressive
 CC supranuclear palsy (PSP), amyotrophic lateral sclerosis, Huntington's

CC disease, multiple sclerosis, and other degenerative diseases such as
 CC cardiovascular disease and rheumatoid arthritis.

CC Sequence 386 AA;

QY Query Match 1.4%; Score 117; DB 22; Length 386;
 QY Best Local Similarity 28.0%; Pred. No. 5.36e+00;

QY Matches 30; Conservative 26; Mismatches 50; Indels 1; Gaps 1;

Db 228 aggplyr-gravrarprdarraaargarsatpgagaparraraparrrrpaarg 286
 QY 232 AGVPLGAPAGARRRGGSASRSLPLPKRRGAPEPERTVGGSGMAHGRTRGSPSD 291

Db 287 grggpragagragrgrgrrgaagagagavrlpaapprrgrga 333
 QY 292 FCVVSPPAPPAEATSLGALSGTRHSHPSVGRQHHAGPSTSRP 338

RESULT 3

ID R23846 standard; Protein; 399 AA.

AC R23846; 05-NOV-1992 (first entry)

DE Prepro form of human osteogenic polypeptide-2.

KW Human osteogenic polypeptide-2; prepro sequence; implant;
 KW endochondral bone formation; skeletal abnormality; dental abnormality;
 KW non-union fracture; cartilage repair; osteoarthritis;

KW periodontal applicn.

KW Homo sapiens.

OS Homo sapiens.

FT Key Location/Qualifiers

FT cleavage_site 13..18 /label= cleavage site for signal peptide removal

FT cleavage_site 257..261 /label= cleavage site to form mature protein

FT protein 261..399 /label= mature protein 'hOP2-A'

FT protein 264..399 /label= mature protein 'hOP2-A'

FT protein 267..399 /label= truncated protein 'hOP2-P'

FT protein 267..399 /label= truncated protein 'hOP2-R'

FT protein 240..399 /label= truncated protein 'hOP2-S'

FT protein 303..399 /label= conserved six cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

FT region 297..399 /label= conserved seven cysteine skeleton

KW nephropathy; glomerulopathy; hypertrophy; sclerosis; nephritis; mouse;
 KW dysplasia; fibrosis; glomerular filtration rate; GFR.
 OS Mus sp.
 FH Key Location/Qualifiers
 FT Peptide 1..17
 FT Peptide /label= signal
 FT Peptide 18..263
 FT Protein /label= propeptide
 FT Protein 264..402
 FT Protein /label= Op-2
 PN W09741881-A1.
 PD 13-NOV-1997.
 PE 06-MAY-1997; U07816.
 PR 06-MAY-1996; US-643321.
 PA (CREA-) CREATIVE BIOMOLECULES INC.
 PI Cohen CM, Sampath KT;
 DR WPI: 97-558690/51.
 DR N-PSDB: V10348.
 PT Treatment of chronic renal failure - using an osteogenic
 PT protein/bone morphogenetic protein renal therapeutic agent or
 PT morphogen or renal mesenchymal progenitor cells
 PS Disclosure: Page 68-69; 113pp; English.
 CC This sequence represents the mouse osteogenic protein, OP-2. This
 CC protein is used in a method for the treatment of a mammal having, or at
 CC risk of, chronic renal failure which comprises administering an
 CC osteogenic protein/bone morphogenetic protein (OP/BMP) renal therapeutic
 CC agent or morphogen. The method can be used for treating e.g. chronic
 CC renal failure, end-stage renal disease, chronic diabetic nephropathy,
 CC diabetic glomerulopathy, diabetic renal hypertrophy, hypertensive
 CC nephrosclerosis, hypertensive glomerulosclerosis, chronic
 CC glomerulonephritis, hereditary nephritis, renal dysplasia, or a patient
 CC afflicted with e.g. glomerular hypertrophy, tubular hypertrophy,
 CC glomerulosclerosis, tubulointerstitial sclerosis or renal fibrosis. Such
 CC therapeutic agents can prevent, inhibit or delay the progressive loss of
 CC functional nephron units and the progressive decline in glomerular
 CC filtration rate (GFR) which slowly but inevitably leads to the need for
 CC renal replacement therapy.
 SO Sequence 399 AA;

Query Match 1.4%; Score 118; DB 29; Length 399;
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

DB 12 glalcalgsgghprpp-htc-pqrrlg-arerrdmgrlavlglpgr-prprgpaar 67
 QY 194 GPPLTQLGAATQARPPHSGPRRLGGERAMNHSVREAGVPLGAPAGARRGGSASRS 253
 DB 68 qp 69
 QY 254 LP 255

RESULT 5
 ID W44314 standard; Protein; 399 AA.
 AC W44314;
 DT 27-MAY-1998 (first entry)
 DE Mouse osteogenic protein OP2.
 KW Mouse; osteogenic protein; subunit; endochondral bone formation;
 KW dimeric; recombinant protein.
 OS Mus sp.
 PN US5714589-A.
 PD 03-FEB-1998.
 PE 23-MAY-1995; 447570.
 PR 21-FEB-1992; US-841646.
 PR 08-APR-1988; US-179406.
 PR 15-AUG-1988; US-232630.
 PR 23-FEB-1989; US-315342.
 PR 17-OCT-1989; US-422613.
 PR 17-OCT-1989; US-422699.
 PR 22-FEB-1990; US-483913.
 PR 20-AUG-1990; US-569920.
 PR 07-SEP-1990; US-579665.
 PR 18-OCT-1990; US-599543.

PR 18-OCT-1990; US-600024.
 PR 21-NOV-1990; US-616374.
 PR 04-DEC-1990; US-621849.
 PR 04-DEC-1990; US-621988.
 PR 22-FEB-1991; US-660162.
 PR 20-DEC-1991; US-810560.
 PR 28-JAN-1992; US-827052.
 PR 01-NOV-1993; US-147023.
 PR 23-MAY-1995; US-447570.
 PA (STYC) STRYKER CORP.
 PI Kuderampath T, Oppermann H, Ozkaynak E, Pang RHL,
 PI Rueger DC;
 DR WPI: 98-158353/14.
 DR N-PSDB: V15217.
 PT Extraction of osteogenic protein from mixture - using antibodies
 PT specific for novel polypeptide chains useful as subunit(s) of
 PT dimeric osteogenic protein(s)
 PS Disclosure: Column 137-140; 127pp; English.
 CC The present sequence represents a mouse osteogenic protein, which is
 CC used in the present invention. The present invention describes methods
 CC for selectively extracting an osteogenic protein (OP) from a mixture.
 CC The method comprises: (a) exposing the mixture to an antibody that
 CC specifically binds OP, separating the resulting antibody-protein complex
 CC from the mixture, and dissociating the complex. In the methods OP
 CC comprises a pair of oxidised subunits that are disulphide-bonded to form
 CC a dimer, and one of the subunits has an amino acid (aa) sequence
 CC sufficiently homologous to residues 335-431 of a 431 aa protein
 CC designated OP5, sequence given in the specification. In dimeric form OP
 CC is capable of inducing cartilage and endochondral bone formation in a
 CC mammal when disposed within a matrix implanted in the mammal. The
 CC methods are used for recovering the recombinant proteins from cell
 CC cultures.
 SO Sequence 399 AA;

Query Match 1.4%; Score 118; DB 29; Length 399;
 Best Local Similarity 38.7%; Pred. No. 4.60e+00;
 Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

DB 12 glalcalgsgghprpp-htc-pqrrlg-arerrdmgrlavlglpgr-prprgpaar 67
 QY 194 GPPLTQLGAATQARPPHSGPRRLGGERAMNHSVREAGVPLGAPAGARRGGSASRS 253
 DB 68 qp 69
 QY 254 LP 255

RESULT 6
 ID R85759 standard; Protein; 399 AA.
 AC R85759;
 DT 19-JUN-1996 (first entry)
 DE MOP-2.
 KW Human; osteogenic protein; hOP-1; murine; MOP-1; TGF-beta superfamily;
 KW transforming growth factor-beta; dimer; antibody; epiloque; hippocampus;
 KW purification; implantable osteogenic device; bone formation; craniofacial;
 KW anomaly; skeletal; dental; endochondral bone formation;
 KW non-union fracture; cartilage repair; osteoarthritis.
 OS Mus musculus.
 FH Key Location/Qualifiers
 FT Peptide 1..260
 FT Peptide /note= "Prepro-peptide"
 FT Protein 261..399
 FT Protein /note= "Mature MOP-2, also designated MOP2-ALa,
 FT Peptide claim 5"
 FT Peptide 303..399
 FT Peptide /note= "MOP-2 peptide, claim 5"
 FT Peptide 17..260
 FT Peptide /note= "MOP-2 peptide, claim 5"
 PN US5468845-A.
 PD 21-NOV-1995.
 PE 08-APR-1988; 179406.
 PR 08-APR-1988; US-179406.
 PR 15-AUG-1988; US-232630.

Query Match	1.4%	Score 118:	DB 7:	Length 399:
Best Local Similarity	38.7%	Pred. No. 4.60e+00:		
Matches 24:	Conservative 12:	Mismatches 22:	Indels 4:	Gaps 4:
Db	68 qp 69			
QY	254 LP 255			
<p>RESULT 9</p> <p>ID R46744 standard; Protein; 399 AA.</p> <p>AC R46744;</p> <p>DT 25-AUG-1994 (first entry)</p> <p>DE Mouse osteogenic protein mop2.</p> <p>KW mouse osteogenic protein; mop2; murine; morphogen;</p> <p>KM infant food formulation; tissue morphogenesis; tissue development;</p> <p>KV bone growth; morphogen-enriched nutritional product.</p> <p>OS Muridae.</p> <p>PN MO9403075-A.</p> <p>PD 17-FEB-1994.</p> <p>PE 29-JUL-1993: U07190.</p> <p>PF 31-JUL-1992: US-923780.</p> <p>PR 31-JUL-1992: US-922813.</p> <p>PR 16-SEP-1992: US-946235.</p> <p>PR 04-MAR-1993: US-029335.</p> <p>PR 31-MAR-1993: US-040510.</p> <p>PA (CREA-) CREATIVE BIOMOLECULES INC.</p> <p>PI Jones WK, Ruberasampath T, Oppermann H, Ozkaynak E;</p> <p>PI Rieger DC, Tucker RF, Cohen CM, Pang RHL;</p> <p>PI WPI: 94-065304/08.</p> <p>DR N-PSDS; Q58052.</p> <p>PT Morphogen enriched dietary compositions and infant formula - capable of enhancing tissue morphogenesis, development and viability, e.g. in infants, aged individuals and metabolic disorders, e.g. anorexia nervosa, etc</p> <p>PS Disclosure and Claims 25-26; Page 127-129; 160pp; English.</p> <p>CC Murine osteogenic protein mop2 and proteins having at least 70% homology with it are preferred morphogens for inclusion in new morphogen-enriched nutritional formulations. The formulations are dietary compositions suitable for people at risk for tissue damage due to protein energy malnutrition or to altered metabolism function and infant formulations to enhance tissue development in an infant or juvenile.</p> <p>CC Sequence 399 AA;</p> <p>SO</p>				
Query Match	1.4%	Score 118:	DB 9:	Length 399:
Best Local Similarity	38.7%	Pred. No. 4.60e+00:		
Matches 24:	Conservative 12:	Mismatches 22:	Indels 4:	Gaps 4:
Db	12 g1a1calagghppp-htc-pqrrlg-arecidmqreila1v1gpr-prrpaqaar 67			
QY	194 GPPLYQGAAGQARPPPHASGPRRLGGERAMNHSVRAGVPLDLPAGARRGGSASRS 253			
Db	68 qp 69			
QY	254 LP 255			
<p>RESULT 10</p> <p>ID R50201 standard; Protein; 399 AA.</p> <p>AC R50201;</p> <p>DT 11-OCT-1994 (first entry)</p> <p>DE Murine OP-2.</p> <p>KW OP-1; OP-2; CBMP2; Vg1(fx); Vgr(fx); DPP(fx);</p> <p>KM GDE-1(fx); 60A(fx); BMP3(fx); BMP6(fx);</p> <p>KM osteogenic protein; morphogen; morphogenic protein;</p> <p>KM gastrointestinal tract; luminal lining; epithelial cell;</p> <p>KM proliferation; ulcer; lesion; inflammation; regeneration;</p>				

OS	Mus musculus.
KW	tissue.
PN	WP09406420-A.
PD	31-MAR-1994.
PF	15-SEP-1993; U08885.
PR	15-SEP-1992; US-945286.
PR	04-MAR-1993; US-029335.
PR	31-MAR-1993; US-040510.
PA	(CREA-) CREATIVE BIOMOLECULES INC.
P1	Charlette WF, Cohen CM, Kuderampath T, Oppermann H;
P1	Ozkaynak E, Pang RHL, Rueger DC, Smart JE;
DR	N-PDSB: 045119.
PT	Maintaining integrity of gastrointestinal lining using a
PT	morphogen (stimulant) - for treating or preventing ulceration,
PT	also to inhibit endothelial cell proliferation and reduce side
PT	effects of cancer therapy.
PS	Claim 35-36; Page 115-117; 151pp; English.
CC	Morphogens comprising an amino acid sequence sharing at least
CC	70% homology with OP-1, OP-2, BMP2, BMP3(fx), Vgl(fx), Vgr(fx),
CC	BMP(fx), GDF-1(fx), 60d(fx) and at least 80% homology with
CC	DMP(fx) and BMD6(fx) are useful for maintaining the integrity of
CC	the gastrointestinal tract luminal lining in a mammal, including
CC	(1) limiting epithelial cell proliferation, (2) inhibiting ulcerative
CC	lesion formation, (3) inhibiting inflammation normally associated
CC	with ulcerative diseases, and/or (4) stimulating the repair of
CC	ulcerative lesions and the regeneration of the luminal tissue.
SQ	Sequence 399 AA;
Query March	1.4%; Score 118; DB 10; Length 399;
Best Local Similarity	38.7%; Pred. No. 4,60e+00;
Matches 24; Conservative	12; Mismatches 22; Indels 4; Gaps 4;
Db	12 g1acalcagghqppr-htc-pqr1g-aretdmqreilaivlgpr-dprpqpaar 67
Oy	194 GPPLYGGAQAAGARPPIASGPERRRLCGERAMNNSVREAGVPGLPAPGARRGGSASRS 253
Db	68 qp 69
Oy	254 LP 255
RESULT	11
ID	R47253 standard; Protein; 399 AA.
AC	R47253;
DT	15-AUG-1994 (first entry)
DE	MOP2.
KW	Human: hippocampus; osteogenic protein; OP-1; subunit; dimer;
KW	morphogenic activity; cysteine; morphogen; family; pro-region;
KW	complex; soluble; aqueous solvent; therapeutic composition;
KW	symptom-alleviating; co-factor; antibody; diagnosis; assay;
KW	quantitate; mature.
OS	Mus musculus.
PN	WP09403600-A.
PD	17-FEB-1994.
PF	29-JUL-1993; U07189.
PR	31-JUL-1992; US-923780.
PR	04-MAR-1993; US-029335.
PR	31-MAR-1993; US-040510.
PA	(CREA-) CREATIVE BIOMOLECULES INC.
P1	Jones WK, Kuderampath T, Oppermann H, Ozkaynak E;
P1	Rueger DC, Tucker RF;
DR	WP1: 94-065689/08.
DR	N-PDSB: 056201.
PT	Morphogenic protein soluble complex - for regeneration of tissue
PT	in mammals and diagnosing tissue disorders
PS	Claim 3; Page 75-77; 120pp; English.
CC	This sequence represents the murine derived protein, osteogenic
CC	protein, MOP-2. The mature OP-2 protein was used as at
CC	least one subunit in the dimeric protein of the invention. This
CC	dimeric protein comprises a pair of protein subunits which are
CC	associated to give a structure with morphogenic activity. Each
CC	subunit comprises more than 100 amino acids having a pattern of

RESULT 14
ID R57974 standard: Protein; 399 AA.
AC R57974: (first entry)
DE Murine OP-2.
OP-1: OP-2; CBMP2; Vg1(fx); Vgr(fx); DPP(fx);
KM GPF-1(fx); 60A(fx); BMP3(fx); BMP5(fx); BMP6(fx);
KM tooth socket; alveolus; osteogenic protein; morphogen;
KM morphogenic protein; periodontal tissue; regeneration;
KM tooth implant; integration; inhibition.
OS Mus musculus.
PN WO9406399-A.
PD 31-MAR-1994.
PE 15-SEP-1993: U08742.
PR 15-SEP-1992: US-945285.
PR 04-MAR-1993: US-029335.
PR 31-MAR-1993: US-040510.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PI Cohen CM, Kuberampath T, Oppermann H, Ozkaynak E;
PI Pang RHL, Rueger DC, Smart JE;
DR WPI: 94-118107/14.
DR N-PSDB: Q67314.
PT Morphogen-induced periodontal tissue regeneration - used in
PT integrating as implanted tooth in tooth socket or to inhibit
PT tissue loss associated with periodontal disease or injury
PS Claim 28-29: Page 100-103: 132pp: English.
CC Morphogens comprising an amino acid sequence sharing at least
CC 70% homology with OP-1, OP-2, CBMP2, Vg1(fx), Vgr(fx), DPP(fx),
CC GPF-1(fx), 60A(fx) and at least 80% homology with BMP3(fx),
CC BMP5(fx) and BMP6(fx) are useful for integrating an implanted
CC tooth in a tooth socket and for inhibiting tissue loss associated
CC with periodontal disease or injury.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 10; Length 399;
Best Local Similarity 38.7%; Pred. No. 4.60e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalagggghppp-htc-pqrrlg-aretdmqrellavlgpgr-prpraqpaar 67
QY 194 GPPLYQLGAATQARPPPHASGPRRLRCERAMNHSVREAGVPLGLPAPGARRGGSASRS 253
Db 68 qp 69
QY 254 LP 255

RESULT 15
ID R33411 standard: Protein; 399 AA.
AC R33411:
DE 15-JUL-1993 (first entry)
DE Murine Pro-OP-2-PP.
KM morphogenic; osteogenic protein; developmental cascade; mop-2;
KM mouse; inflammation; anti-inflammatory; Transforming Growth Factor;
KM TGF-beta super-family; hippocampus.
OS Mus.
FH Key Location/Qualifiers
FT cds 93..1292 /*tag- a
FT mat_peptide 873..1289 /note- "mop-2-pp"
FT /*tag- b
PN WO9304692-A. /note- "contains conserved 7 cysteine skeleton"
PD 18-MAR-1993.
PE 28-AUG-1992: U07358.
PR 30-AUG-1991: US-752764.
PR 30-AUG-1991: US-752861.
PR 30-AUG-1991: US-753059.
PA (CREA-) CREATIVE BIOMOLECULES INC.
PI Cohen CM, Kuberampath T, Oppermann H, Ozkaynak E;
PI Pang RHL, Rueger DC, Smart JE;
DR WPI: 93-100652/12.

DR N-PSDB: Q38736.
PT Morphogen-induced modulation of inflammatory response - and
PT resulting tissue damage, e.g. in autoimmune diseases, diabetes,
PT asthma, ischemia reperfusion injury, etc.
PS Claim 26: Page 126-127: 165pp: English.
CC Osteogenic protein (OP)-2 is a preferred morphogen
CC for use in treating tissue damage in e.g. inflammatory disease,
CC autoimmune disease, arthritis, psoriasis, dermatitis, diabetes and
CC emphysema. Proteins having at least 70% homology with OP-2 amino
CC acid sequences can also be used. See R33401 for mature mop-2.
SQ Sequence 399 AA;

Query Match 1.4%; Score 118; DB 7; Length 399;
Best Local Similarity 38.7%; Pred. No. 4.60e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 glalcalagggghppp-htc-pqrrlg-aretdmqrellavlgpgr-prpraqpaar 67
QY 194 GPPLYQLGAATQARPPPHASGPRRLRCERAMNHSVREAGVPLGLPAPGARRGGSASRS 253
Db 68 qp 69
QY 254 LP 255

Search completed: Fri Dec 18 18:25:06 1998
Job time : 125 secs.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Db 281 APPAGKPGPPPPGGRPPR 302
QY 317 SHPSVGRHHAAGPSTGRPPR 338

RESULT 7
ENTRY S16681 #type complete
TITLE homeotic protein - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 17-Oct-1997

ACCESSIONS
REFERENCE S16681
#authors Deguchi, Y.; Kehrl, J.H.
#journal Nucleic Acids Res. (1991) 19:3742
#title Nucleotide sequence of a novel diverged human homeobox gene encodes a DNA binding protein.
#cross-references MUID:91305125
#accession S16681
#status preliminary; nucleic acid sequence not shown;
translation not shown

##molecule_type mRNA
##residues 1-316 #label DEG
##cross-references EMBL:X56537
##note the nucleotide sequence was submitted to the EMBL Data Library, January 1991

CLASSIFICATION #superfamily unassigned homeobox proteins; homeobox homology
KEYWORDS DNA binding; homeobox; nucleus; transcription regulation
FEATURE
73-127 #domain homeobox homology #label HOX
SUMMARY #length 316 #molecular-weight 34713 #checksum 3095

Query Match 1.6%; Score 136; DB 2; Length 316;
Best Local Similarity 30.9%; Pred. No. 1.87e-03;
Matches 21; Conservative 20; Mismatches 24; Indels 3; Gaps 3;

Db 202 RPNWTSASDCSVGRGAPAR-SPPRSRPAQDSRPARSPGPAASGGPGAMTH 260
QY 223 RAMNHSVREACV-PLGLPAGARRRGSGASRLPLPKRRG-AAPEPERTVGGGWAH 280

Db 261 PAPPREQA 268
QY 281 PGRTRGPS 288

RESULT 8
ENTRY S01955 #type complete
TITLE hypothetical protein, 69K - turnip yellow mosaic virus
ORGANISM #formal_name turnip yellow mosaic virus; TYMV
DATE 21-Nov-1993 #sequence_revision 26-May-1995 #text_change 26-May-1995

ACCESSIONS
REFERENCE S01955
#authors Morch, M.D.; Boyer, J.C.; Haenni, A.L.
#journal Nucleic Acids Res. (1988) 16:6157-6173
#title Overlapping open reading frames revealed by complete nucleotide sequencing of turnip yellow mosaic virus genomic RNA.
#cross-references MUID:88289359
#accession S01955
#status preliminary
##molecule_type genomic RNA
##residues 1-628 #label MOR
##cross-references EMBL:X07441
##note the authors translated the codon ACG for residue 459 as U

SUMMARY #length 628 #molecular-weight 69194 #checksum 7569

Query Match 1.6%; Score 137; DB 2; Length 628;
Best Local Similarity 25.0%; Pred. No. 1.40e-03;
Matches 73; Conservative 64; Mismatches 134; Indels 21; Gaps 20;

Db 157 GPVLTFRPRSVQPRSATRGPSRPLLRKV-VH-VHDDPHSLRPRRSRNLQPT 214

QY 194 GPPLYOUGAATQARPPPHAS-GPR-RLGGERAMNHSVREACVPLGLPAGAR-RR-GGS 249
Db 215 VRRPLAPNQHSRPPDISDDGILGPRPLAPHSTRDPPRPI-TPGPSN-THDLRL 272
QY 250 ASRSLPLPKRRRGAAPERTPVG-OGSWA-HPGTRGSDRGFCVSPARPAEBAETS 307
Db 273 SVLPRTSPRRGLPNRRHNTSTGHIPPTTSRPTGPPSLQRPVH-LYQSSPHTNFR 331
QY 308 EG-ALSTGRSHSVGQHAGPPSTSRPPPDTCPPVYATKHFV-SSGDKEQLAR 365

Db 332 SSIRKDALQTPRLGHELEGPANLTERSESPPTKRRLPREESEPNRLPKPLEATLAP 391
QY 366 SFLL-SSL-R-PSLIGARLVETIFLGSPPMPPGTPRRLPRPQRYXWQRRPLELLG- 420

Db 392 SYRRRPPPLPNPPALPSIAVTSSRGKIHSPLKCALPK-EGAPPPRL 442
QY 421 NHAQC-PYGVLTAKHCPRLA-AVTPAAGVCAREKPGGSVAPEEDTDPRRL 470

RESULT 9
ENTRY D34768 #type complete
TITLE ORF4 protein - Orf virus (strain Nz2)
ORGANISM #formal_name Orf virus
DATE 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change 31-Oct-1997

ACCESSIONS
REFERENCE D34768
#authors Fraser, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.
#journal Virology (1990) 176:379-389
#title Sequence analysis of the inverted terminal repetition in the genome of the parapoxvirus, orf virus.
#cross-references MUID:90266454
#accession D34768
#status preliminary
##molecule_type DNA
##residues 1-264 #label FRA
##cross-references GB:M30023; EMBL:M37623; NID:g332561; PID:g332566
SUMMARY #length 264 #molecular-weight 25613 #checksum 1957

Query Match 1.5%; Score 129; DB 2; Length 264;
Best Local Similarity 28.2%; Pred. No. 1.39e-02;
Matches 33; Conservative 40; Mismatches 37; Indels 7; Gaps 5;

Db 150 PAPERAGRYPAAPAAASARRHRCPPAARSPAAERPAAPARAEAPRASSS-GS 208
QY 223 RAMNHSVR-EAGVPLGLPAGARRRGSGASRLPLPKRP--R-RGAAPERTPVGGGS 277

Db 209 RASSGPRRSAAAS-SAAGSRPAAGSAPARPAASARSAGEAARGAPRAGM 264
QY 278 WAHPGRTGPSDGRFCVSPARPAEATSLGALSGTRHSHPSVGRHHAAGPSTSR 334

RESULT 10
ENTRY S19560 #type complete
TITLE proline-rich protein MP4 - mouse
ORGANISM #formal_name Mus musculus #common_name house mouse
DATE 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Mar-1998

ACCESSIONS
REFERENCE S19560; S22570
#authors Roberts, S.G.E.; Layfield, R.; Bannister, A.J.; McDonald, C.J.
#journal Eur. J. Biochem. (1991) 202:969-974
#title Gene sequence of mouse B-type proline-rich protein MP4. Transcriptional start point and an upstream phylogenetic footprint with ets-like and rel/NKX-like elements.
#cross-references MUID:92111548
#accession S19560
#status preliminary
##molecule_type DNA
##residues 1-300 #label ROB
##cross-references GB:X58438; NID:g53181; PID:g53182

```

REFERENCE          S22570
#authors          Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
#journal          Nucleic Acids Res. (1991) 19:5205-5211
#title            The mouse proline-rich protein Mp promoter binds
                  isoprenolene-inducible parotid nuclear proteins via a
                  highly conserved NFkB/rel-like site.
#cross-references MIMD:92020206
#accession        S22570
#molecule_type  DNA
#residues          1-14 ##label RO2
#cross-references EMBL:X61126
CLASSIFICATION    #superfamily proline-rich protein
SUMMARY           #length 300 #molecular-weight 31129 #checksum 8443

Query Match
Best Local Similarity 25.2%; Pred. No. 7,89e-03;
Matches 34; Conservative 40; Mismatches 54; Indels 7; Gaps 7;

Db 131 PGNOGPPPOG-GPQQRRP-QPGNOGPPPPGGGQRRPPPOGGGNOGPP-OGGPHP-PPR 186
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 202 AATQRRPPPHASGPPRRRIGCERAMNHVSREGVPLGLPAPGARRRGGSASSLPLPRPR 261
      ::||::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db 187 PGNOGPPPOGQPOORPQPGNOGPPPOGQ6-POAPRRPGNOGPPPOGPPPTGNOQ 245
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 262 RGAPEPRRTYGGGSMHFGRTKPSDRCFCVVSPPARPA-EEATSLDGALSGTRHSHPS 320
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
Db 246 -GPPPGGGPQGGPPRP 259
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 321 VGRQHMGAPPRSTRP 335
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||

RESULT 11
ENTRY   S11790 #type complete
TITLE   nodulation protein nodd3 - Rhizobium leguminosarum bv.
        phaseoli
ORGANISM
        #formal name Rhizobium leguminosarum bv. phaseoli
        #1-Nov-1993 #sequence revision 13-Jan-1993 #text change
        16-Feb-1997
DATE
ACCESSIONS S11790
REFERENCE S11786
#authors Davis, E.O.; Johnston, A.W.B.
#journal Mol. Microbiol. (1990) 4:921-932
#title Analysis of three nod genes in Rhizobium leguminosarum
        biovar phaseoli; nodd1 is preceded by nlpE, a gene whose
        product is secreted from the cytoplasm.
#cross-references MIMD:91014692
#accession S11790
#status not compared with conceptual translation
#molecule_type DNA
#residues 1-302 ##label DAY
GENETICS
#gene nodd3
CLASSIFICATION #superfamily regulatory protein lysR
KEYWORDS DNA binding; transcription regulation
SUMMARY #length 302 #molecular-weight 34274 #checksum 3819

Query Match
Best Local Similarity 31.4%; Pred. No. 3,23e-02;
Matches 38; Conservative 28; Mismatches 46; Indels 9; Gaps 9;

Db 17 LMIEENLTAAARSINLSOPAMSAAVRL-RSYFDEL-FTMRGGEFVTP-R-AEDLAPA 72
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 9 LLRTWEALAPATP-AMRAPRCRAVRSLSRSHYREVLPPLAFVVRLLPGQGRVLQGDPA 67
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||

Db 73 IREALQHRNLNIPMDKTPDOSDRHFR-VS-ICDFVTVVYFQKILELAEAGISIDL 130
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 68 AFRAALVAOCLVCPWDA-RPPAPSRFQVSCLEKELVARVL-ORLCERGAAVNLAFGAL 125
      |::||::||::||::||::||::||::||::||::||::||::||::||::||::||
OY 126 L 126

```

RESULT	12	
ENTRY	S10889	#type complete
TITLE	proline-rich protein - human	
ORGANISM	#format_name Homo sapiens #common_name man	
DATE	07-oct-1994	#sequence_revision 26-May-1995
	08-Sep-1997	#text_change
ACCESSIONS	S10889	
REFERENCE	S02127	
#authors	Lyons, K.M.; Stein, J.H.; Smithies, O.	
#journal	Genetics (1988) 120:267-278	
#title	Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over.	
#cross-references	MUTD:89121440	
#accession	S10889	
#status	preliminary; translation not shown	
#molecule_type	DNA	
##residues	1-309 ##label LYX	
GENETICS	##cross-references EMBL:X07881; NID:g35637; PID:g296669	
CLASSIFICATION	22/1; 34/1	
#introns	#superfamily proline-rich protein	
SUMMARY	#length 309 #molecular_weight 30936 #checksum 3043	
	Query Match 1.5%; Score 129; DB 2; Length 309;	
	Best Local Similarity 30.4%; Pred.No. 1.3e-02;	
Matches	45; Conservative 38; Mismatches 53; Indels 12; Gaps	
Dd	122 GPPP-QGGNSQG-PPPHPRKPEBPPP-QGQ-NOSGPPRRP-GKE-GEPPPGGNSQG	175
	::: ::: ::: ::: :::	
Qy	194 GPPLYQLGAQAAPPBPASGPBRRLCERAMNHSVREAGVPDLPAAPGARRRGSASRS	253
	::: ::: ::: ::: :::	
Dd	176 -P-PPRPCKRPGPPPGGNSQSGPPPRPKREDGPPPGGNSQSGPPPRPKREGSPSGG	233
	::: ::: ::: ::: :::	
Qy	254 LPLEPRRPRGAAPPERTPVYQSGWAPRGRTGPSDRG-F-CVVSAPRAE-EXTSLIGA	310
	::: ::: ::: ::: :::	
Dd	234 -NKRPGRPHPKPGPPPGGNSQSGPPRRPKRPPR 260	
	::: ::: ::: ::: :::	
Qy	311 LSGTRHSHPSVGROHNAGPSTSRPPRP 338	
	::: ::: ::: ::: :::	

```

RESULT      13
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
#authors
#journal
#title

#cross-references MUID:88273214
#accession      A28996
##molecule_type DNA
##residues      1-317 ##label ANN
#cross-references GB:M2326; GB:J03891; NID:9200535; PID:9567232

GENETICS
#introns
CLASSIFICATION
KEYWORDS
FEATURE
1-15
16-317

SUMMARY
#length 317 #molecular-weight 31719 #checksum 8454

Query Match      1.5%: Score 132; DB 2; Length 317;
Best Local Similarity 26.4%: Pred. No. 5, 93e-03;
Matches 39; Conservative 38; Mismatches 64; Indels 7; Gaps 6;
37 GSGPPRPVNGSGDGGPPGPGQPRD-PQGGPPGPGQPRPQGGPPGPGQPRD-PQGG

```

WU

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:25:25 1998; MasPar time 30.89 Seconds
Tabular output not generated. 1002.720 Million cell updates/sec

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect Score: 8624
Sequence: 1 HSGGQRCVLTWELALAPAT.....TALEAANPALPSDFITLD 1154

Scoring table: PAM 150
Gap 11

Searched: 74019 segs, 26840295 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot35
1:swissprot

Statistics: Mean 55.898; Variance 96.260; scale 0.581

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	162	1.9	660	1	YH1_EBV	4.64e-09
2	153	1.8	504	1	ATIN_HSVBP	1.47e-07
3	150	1.7	234	1	ALPHA TRANS-INDUCING P	4.56e-07
4	142	1.6	276	1	PRPM_HUMAN	8.72e-06
5	137	1.6	628	1	SALIVARY PROLINE-RICH	5.26e-05
6	133	1.5	247	1	V70K_TYRV	2.15e-04
7	131	1.5	296	1	PRP4_HUMAN	4.31e-04
8	127	1.5	302	1	PROLINE-RICH PROTEIN M	2.37e-03
9	126	1.5	566	1	TESTIS-SPECIFIC PROTEIN D	1.69e-03
10	125	1.4	261	1	TS13_MOUSE	3.32e-03
11	121	1.4	322	1	PRP2_MOUSE	1.25e-02
12	118	1.4	399	1	PROLINE-RICH PROTEIN D I	3.30e-02
13	115	1.4	431	1	BONE MORPHOGENETIC PRO	3.32e-03
14	118	1.4	401	1	BONE MORPHOGENETIC PRO	3.30e-02
15	125	1.4	501	1	MESENCHYME FORK HEAD P	3.32e-03
16	118	1.4	514	1	REGULATORY PROTEIN E2	3.30e-02
17	114	1.3	202	1	VE2_HPV05	1.17e-01
18	116	1.3	270	1	HYPOHETICAL GENE 1 PR	6.23e-02
19	115	1.3	279	1	TRNA-(MSI2)10(6)A-KD	8.54e-02
20	109	1.3	301	1	HYPOHETICAL 29.3 KTD	5.40e-01
21	114	1.3	321	1	MODULATION PROTEIN D3	1.17e-01
22	109	1.3	324	1	MODULATION PROTEIN D1	5.40e-01
23	115	1.3	331	1	MATD_NEUCR	8.54e-02

24	114	1.3	408	1	CGE1_XENLA	61/S-SPECIFIC CYCLIN E	1.17e-01
25	114	1.3	408	1	CGE2_XENLA	G1/S-SPECIFIC CYCLIN E	1.17e-01
26	110	1.3	408	1	VE2_XENLA	G1/S-SPECIFIC CYCLIN E	3.99e-01
27	111	1.3	498	1	VE2_HP08	REGULATORY PROTEIN E2	2.95e-01
28	111	1.3	509	1	VE2_HP36	REGULATORY PROTEIN E2	2.95e-01
29	113	1.3	514	1	VE2_HP5B	REGULATORY PROTEIN E2	1.59e-01
30	110	1.3	628	1	VE2_HP5B	REGULATORY PROTEIN E2	3.99e-01
31	112	1.3	633	1	LA17_YEAST	69 KD PROTEIN	2.17e-01
32	113	1.3	704	1	SYN1_RAT	SYNAPSINS IA AND IB	1.59e-01
33	114	1.3	705	1	SYN1_HUMAN	SYNAPSINS IA AND IB	1.17e-01
34	108	1.3	706	1	SYN1_BOVIN	SYNAPSINS IA AND IB	7.28e-01
35	113	1.3	1255	1	MDC1_HUMAN	MUCIN 1 PRECURSOR (POL	1.59e-01
36	115	1.3	3149	1	TEGU_EBV	LARGE TEGUMENT PROTEIN	8.54e-02
37	107	1.2	174	1	BAR1_CHITE	BALBIANT RING PROTEIN	9.79e-01
38	106	1.2	233	1	YFR1_ECOLI	HYPOHETICAL TRANSCRIP	1.31e+00
39	106	1.2	285	1	YAF1_ECOLI	HYPOHETICAL TRANSCRIP	1.31e+00
40	107	1.2	318	1	MODD_RHILT	MODULATION PROTEIN D	9.79e-01
41	107	1.2	494	1	MESENCHYME FORK HEAD P	EARLY E2A DNA-BINDING	9.79e-01
42	107	1.2	529	1	DNB2_ADE05	PHOSPHORIBOSYL-AMP CYC	1.76e+00
43	105	1.2	863	1	HIS2_NEUCR	ACONITATE HYDRATASE 1	1.76e+00
44	105	1.2	890	1	ACOL1_ECOLI	ACONITATE HYDRATASE 1	1.76e+00
45	105	1.2	2774	1	MAPA_RAT	MICROTUBULE-ASSOCIATED	1.76e+00

ALIGNMENTS

RESULT	ID	YH1_EBV	STANDARD:	PRT:	660 AA.
AC	P03181	21-JUL-1986 (REL. 01, CREATED)			
DT	21-JUL-1986	(REL. 01, LAST SEQUENCE UPDATE)			
DT	15-JUL-1998	(REL. 36, LAST ANNOTATION UPDATE)			
DE	HYPOHETICAL BHLFI PROTEIN.				
OS	EPSTEIN-BARR VIRUS (STRAIN B95-8) (HUMAN HERPESVIRUS 4).				
OC	HERPESVIRIDAE; GAMMAHERPESVIRINAE.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE: 84270667.				
RA	BAER R., BANKIER A.T., BIGGIN M.D., DEININGER P.L., FARRELL P.J.,				
RA	GIBSON T.J., HATFUL G., HUDSON G.S., SATCHELL S.C., SEGUIN C.,				
RA	TUFENELL P.S., BARRELL B.G.;				
RL	NATURE 310:207-211(1984).				
DR	EMBL: V01555; -; NOT_ANNOTATED_CDS.				
PIR	A03742; Q08E3.				
KW	HYPOHETICAL PROTEIN; EARLY PROTEIN; REPEAT.				
FT	DOMAIN 149 648 4 X 125 AA TANDEM REPEATS.				
FT	REPEAT 149 273 1.				
FT	REPEAT 274 398 2.				
FT	REPEAT 399 523 3.				
FT	REPEAT 524 648 4.				
SQ	SEQUENCE 660 AA: 66244 MW; 372F08C5 CRC32;				
Query Match	1.9%; Score 162; DB 1; Length 660;				
Best Local Similarity	30.8%; Pred. No. 4.64e-09;				
Matches	40; Conservative 34; Mismatches 60; Indels 14; Gaps 12;				
DB	251 GPPPTRSAAAG-RTHRRPGCPRSARNPGCPRTYRR-R-SGAORGHPPPGAGRPSPC 306				
QY	194 GPLLYQLGATQARPPPHASG-PRR-R-GCERANVHSYREAGVPLGLPAPGA-RRRGS 249				
DB	307 TCGRAAPACPTPAAPGGAAPVSGATPHBEGSGADPP-AAARLPREPERLP 364				
QY	250 ARSRLPLPRPRRGAPEPER-TPVGGGWAHPGTRGSDGFCVVSAPRAEATISLE 308				
DB	365 ODIAAQR-C-PAGPPTRSQA-AAQTRH-R-PPGCP 397				
QY	309 GALSGRHSHPVGRVGHNAHGPSTSRPPRMDPCP 344				
RESULT	2				
ID	ATIN_HSVBP	STANDARD:	PRT:	504 AA.	
AC	P30020				
DT	01-APR-1993 (REL. 25, CREATED)				

DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
DE ALPHA TRANS-INDUCING PROTEIN (ALPHA-TIF).
OS BOVINE HERPESVIRUS TYPE 1 (STRAIN PB-2).
CC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93012995.
RA CARPENTER D.E., MISRA V.;
RL GENE 119:259-263(1992).
CC -I- FUNCTION: RESPONSIBLE FOR TRANSCRIPTIONAL ACTIVATION OF IMMEDIATE
EARLY PROMOTERS (ALPHA GENES) (BY SIMILARITY).
CC -I- SIMILARITY: TO OTHER HERPESVIRUSES ALPHA TRANS-INDUCING PROTEIN.
DR EMBL: Z11610; E264419; -.
DR PIR: S24229; S24229.
DR JCI306; JCI306.
KW TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR; DNA-BINDING.
SQ SEQUENCE 504 AA; 54028 MW; 79F42020 CRC32;
Query Match 1.8%; Score 153; DB 1; Length 504;
Best Local Similarity 31.3%; Pred. No. 1.47e-07;
Matches 46; Conservative 33; Mismatches 56; Indels 12; Gaps 11;
Db 351 APEAGGGRGRTGRTGARAATGRLQRCGPRRRAC-CRATP-RORLR--ARGE 406
Qy 208 PPPHSGRRRLGRCERAMNHSVREAGVPLGLPAPARRGGSSASLSPLPKRRGAPE 267
Db 407 PRHTS-GSGAESQ-GRRPGVCRGLMACARSGPARGGPSPVNSGLSLF-ARGSPG 463
Qy 268 PRTVVGQGSMAHPGRTGSDR-GF-CVV-S-PARPAEATSLGALSGTSHSPVGR 323
Db 464 GPACGPRRARGRRRASPAN-FGGT 489
Qy 324 OHNAGPPTSRPPRWDTPCPVYAE 350
RESULT 3
ID PRPM_HUMAN STANDARD; PRT: 234 AA.
AC P10161; P02813;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE M) (CONTAINS: PEPTIDE P-D)
DE (FRAGMENT).
GN PRB4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89121439.
RA LYONS K.M., STEIN J.H., SMITHIES O.;
RL GENETICS 120:255-265(1988).
RN [2]
RP SEQUENCE OF 165-234.
RX MEDLINE: 83186122.
RA SALTOW E., ISEMURA S., SANADA K.;
RL J. BIOCHEM. 93:495-502(1983).
DR EMBL: X07704; E265547; -.
DR PIR: A03295; PIHUSD.
DR PIR: S03175; S03175.
DR HSSP: P19999; ICLG.
DR MIM: 168730; -.
DR MIM: 180990; -.
KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.
FT NON-TER 1
FT CHAIN 165 234 PEPTIDE P-D.
SQ SEQUENCE 234 AA; 23676 MW; ED2D4ADC CRC32;
Query Match 1.7%; Score 150; DB 1; Length 234;
Best Local Similarity 29.3%; Pred. No. 4.56e-07;
Matches 43; Conservative 41; Mismatches 53; Indels 10; Gaps 10;

Db 88 GPP-OGGNSOG-PPHPGKPERPP-OGG-NQSHRPPPP-GKPER-PPPOGNSOG 141
Qy 194 GPPYQGAATQAPPPASPPRRRLCCERAMNHSVREAGVPLGLPAPARRGGSSRS 253
Db 142 -P-PPHKGPPPOGNSRSARSPPGKPOGPOGNNKPOGPPPPGKPOGPPPGN 199
Qy 254 LPLPKRRPAGAAPPERTVVGQSMHPRGTRGSDR-GFCVSPAPPAE-EATSLGAL 311
Db 200 PQOPAPPPAGKPPGPPPPGKPPPP 226
Qy 312 SGTSHSPSVGRQHAGPPTSRPPRP 338
RESULT 4
ID PRPL_HUMAN STANDARD; PRT: 276 AA.
AC P10162; P02813;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE SALIVARY PROLINE-RICH PROTEIN PO (ALLELE K) (CONTAINS: PEPTIDE P-D)
DE (FRAGMENT).
GN PRB4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 89121439.
RA LYONS K.M., STEIN J.H., SMITHIES O.;
RL GENETICS 120:255-265(1988).
RN [2]
RP SEQUENCE OF 207-276.
RX MEDLINE: 83186122.
RA SALTOW E., ISEMURA S., SANADA K.;
RL J. BIOCHEM. 93:495-502(1983).
DR EMBL: X07715; E4806; ALT_SEQ.
DR PIR: A03295; PIHUSD.
DR PIR: S03176; S03176.
DR HSSP: P19999; ICLG.
DR MIM: 168730; -.
DR MIM: 180990; -.
KW REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.
FT NON-TER 1
FT CHAIN 207 276 PEPTIDE P-D.
SQ SEQUENCE 276 AA; 27816 MW; 4838945A CRC32;
Query Match 1.6%; Score 142; DB 1; Length 276;
Best Local Similarity 28.2%; Pred. No. 8.72e-06;
Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;
Db 134 QGGNSOG-PPHPGKPERPP-OGG-NQSHRPPPP-GKPER-PPPOGNSOG-P-PP 186
Qy 199 QLGATQARPPHSGRRRLGRCERAMNHSVREAGVPLGLPAPARRGGSSASLSPLK 258
Db 187 HPKREPPPOGNSRSARSPPGKPOGPOGNNKPOGPPPPGKPPPPGPPGPPQ 246
Qy 259 PPRKAAPERTVVGQSMHPRGTRGSDR-GFCVSPAPPAE-EATSLGALSGTRH 316
Db 247 APPAGKPPGPPPPGKPPPP 268
Qy 317 SHPSVGRQHAGPPTSRPPRP 338
RESULT 5
ID V7OK_TYMV STANDARD; PRT: 628 AA.
AC P10357;
DT 01-MAR-1989 (REL. 10, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)
DE 69 KD PROTEIN.
OS TURNIP YELLOW MOSAIC VIRUS.
CC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; TYMOVIRIDAE.
RN [1]

[illegible][illegible]

SQ SEQUENCE 431 AA; 44309 MM; 232AB9D7 CRC32;

Query Match 1.4%; Score 118; DB 1; Length 431;

Best Local Similarity 32.1%; Pred. No. 3.30e-02;

Matches 27; Conservative 20; Mismatches 31; Indels 6; Gaps 6;

Db 268 RGGGKPPGLSPRA-TDGNRDPG-A-GVAPRRRRMGSSGGGTPGGRPERAAPGARP 324

QY 211 HASGPRRLGCCRANWNSVREAGVPLGLPAPGARRRGSSAS-RSLPLKRRPRRGAPEPE 269

Db 325 TAPDAPGRMDGPDADGAPAGLGAG 348

QY 270 RTPVG-QGSMWAPGRTGPPS-DRG 291

RESULT 15

ID MEH1_HUMAN STANDARD; PRT; 501 AA.

AC Q99958;

DT 15-JUL-1998 (REL. 36, CREATED)

DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)

DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE MESENCHYME FORK HEAD PROTEIN 1 (MEH-1 PROTEIN) (TRANSCRIPTION FACTOR

FKH-14).

GN FKH14 OR MEH1.

OS HOMO SAPIENS (HUMAN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;

CC EUHERIA; PRIMATES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE; 97312712.

RA MIURA N., IIDA K., KAKINUMA H., YANG X.-L., SUGIYAMA T.;

RL GENOMICS 41:489-492(1997).

CC -1- FUNCTION: MIGHT BE INVOLVED IN THE FORMATION OF SPECIAL

MESENCHYMAL TISSUES (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).

CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.

DR EMBL; Y08223; E303016; -.

DR MIM; 602402; -.

DR PROSITE; PS00657; FORK_HEAD_1; 1.

DR PROSITE; PS00658; FORK_HEAD_2; 1.

DR PROSITE; PS00659; FORK_HEAD_3; 1.

KW TRANSCRIPTION REGULATION; DNA-BINDING; NUCLEAR PROTEIN.

FT DNA_BIND 71 162 FORK-HEAD.

FT DOMAIN 163 167 POLY-ARG.

FT DOMAIN 387 396 HIS-RICH.

FT DOMAIN 397 421 ALA/PRO-RICH.

FT DOMAIN 400 408 POLY-PRO.

FT DOMAIN 416 422 POLY-ALA.

SQ SEQUENCE 501 AA; 53719 MM; 99D32EFA CRC32;

Query Match 1.4%; Score 125; DB 1; Length 501;

Best Local Similarity 28.4%; Pred. No. 3.32e-03;

Matches 27; Conservative 28; Mismatches 38; Indels 2; Gaps 2;

Db 367 SPLSLNTLACOGALATGHHHGHHPQAPPPAPQPPQPPGAAAGAAASWYL 426

QY 296 SPARAEPAETSEGLSGTRHSHPSVGRQH-HAGPPSTSRPPRPMDTQCPVYATETKFL 354

Db 427 NMSGDLNHL-PGHTFAAOQQTFFPNVREMFSHRLG 460

QY 355 YSSGDKQLRPSFLSLRPSLTGARLVETFTLG 389

Search completed: Fri Dec 18 18:26:40 1998
Job time : 75 secs.

 WIRENET (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

 Msrch_gp protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Dec 18 18:27:02 1998; MasPar time 60.80 Seconds

Tabular output not generated. 945.090 Million cell updates/sec

Title: >US-08-951-733-20

Description: (1-1154) from US08951733.pep

Perfect Score: 8624

Sequence: 1 HASGORCVLRTWEALAPAT.....TALEAANPALPSDFKTLID 1154

Scoring table: PAM 150

Gap 11

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

sptrembl6
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mmc 8:sp_organalle
 9:sp_phase 10:sp_plant 11:sp_prodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 54.168; Variance 101.474; scale 0.534

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	8465	98.2	1132	4	014746	TELOMERASE REVERSE TRA	0.00e+00
2	8459	98.1	1132	4	014783	TELOMERASE CATALYTIC S	0.00e+00
3	5161	59.8	1122	11	070372	TELOMERASE REVERSE TRA	0.00e+00
4	678	7.9	888	3	013339	TELOMERASE REVERSE TRA	1.37e-110
5	671	7.9	889	3	013338	TELOMERASE REVERSE TRA	3.12e-111
6	439	5.1	1031	5	000939	TELOMERASE SUBUNIT P12	1.97e-60
7	424	4.9	67	11	035432	TELOMERASE CATALYTIC S	2.25e-57
8	314	3.6	884	3	006163	CHROMOSOME XII COSMID	1.71e-35
9	147	1.7	296	14	069118	HYPOHETICAL PROTEIN (7.18e-06
10	143	1.7	574	3	030627	WISKOTT-ALDRICH SYNDRO	2.85e-05
11	142	1.6	234	4	000600	PAROTID 'O' PROTEIN (F	4.02e-05
12	141	1.6	276	2	046612	IS 1222 GENE ORF-A AND	5.65e-05
13	136	1.6	316	4	099076	HOMEOBOX PROTEIN (FRAG	3.03e-04
14	140	1.6	373	2	054226	POLYKETIDE SYNTHASE (F	7.93e-05
15	135	1.6	585	14	041935	HYPOHETICAL 60.2 KD P	4.22e-04
16	136	1.6	897	11	070495	PLENTY-OF-PROLINES-101	3.03e-04
17	129	1.5	264	14	085301	HOMOLOGUE OF RETROVIRA	3.01e-03
18	131	1.5	300	11	061888	PROLINE RICH PROTEIN	1.57e-03
19	129	1.5	309	4	004118	SALIVARY PROLINE-RICH	3.01e-03
20	132	1.5	317	11	062103	PROLINE RICH PROTEIN P	1.14e-03

ALIGNMENTS

RESULT ID	1	PRELIMINARY;	PRT;	1132 AA.
AC	014746			
DT	01-JAN-1998 (TREMBLREL. 05, CREATED)			
DT	01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)			
DT	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)			
DE	TELOMERASE REVERSE TRANSCRIPTASE.			
GN	HTRT.			
OS	HOMO SAPIENS (HUMAN).			
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-KIDNEY.			
RX	MEDLINE: 97400623.			
RA	NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,			
RA	LININGER J., HARLEY C.B., CECH T.R.;			
RL	SCIENCE 277:955-959(1997).			
DR	EMBL: AF015950; G2330017; -.			
KW	RNA-DIRECTED DNA POLYMERASE.			
SQ	SEQUENCE 1132 AA; 126995 MM; 2DFBEDF3 CRC32;			
Query Match	98.2%; Score 8465; DB 4; Length 1132;			
Best Local Similarity	100.0%; Pred. No. 0.00e+00;			
Matches	1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
DB	1 MRRARCRVRLSLSHREVLPATFRRRLGPOGWRVQRRDPAAFRALVAQCLVCPW 60			BETA-KETOACYL SYNTHASE
DB	23 MRRARCRVRLSLSHREVLPATFRRRLGPOGWRVQRRDPAAFRALVAQCLVCPW 82			SIMILARITY TO COLLAGEN
DB	61 DRRPPAPRFQVQSCLELVAVRQRLCEKAKNVLAFFGFLDNGARGPPEATTTSVR 120			LF3 PROTEIN.
DB	83 DRRPPAPRFQVQSCLELVAVRQRLCEKAKNVLAFFGFLDNGARGPPEATTTSVR 142			HYPOHETICAL 131.9 KD
DB	121 SYLPPTVTDALRGSGAMGLLRVDDVLYHLARCAFLVAPSCAYOVCPPLXQGA 180			UL36.
DB	143 SYLPPTVTDALRGSGAMGLLRVDDVLYHLARCAFLVAPSCAYOVCPPLXQGA 202			CODED FOR BY C. ELEGAN
DB	181 ATOAAPPPHASPFRRLRCERAMNHSYREAGVPLCLPAPGARRGSGASRSPLPKRRR 240			PROLINE-RICH SALIVARY
DB	203 ATOAAPPPHASPFRRLRCERAMNHSYREAGVPLCLPAPGARRGSGASRSPLPKRRR 262			SUBSTRATE OF THE PROTE
DB	241 GAAPPERTPVGQGWAPRGRTGRGSDRGFCVSPARAEEATSEGLAGSRHSPVSG 300			MPK3030 NODDI GENE IN

```

OY 263 GAAPERPTPYGQSWAHPRGTRGSPDSGFCVVSAPARPAEATSELEGALSGTRSHSPVG 322
Db 301 ROHAGPSTSRPPRPMPTPCPPYAEKHFHYSSGDKQOLRPSFLSSLRSLTGARL 360
OY 323 ROHAGPSTSRPPRPMPTPCPPYAEKHFHYSSGDKQOLRPSFLSSLRSLTGARL 382
Db 361 VETIFGSRPMWPGTPRRLPRLPORYMOMRPLFELLGNHACCPGYLLKTCPLRAVT 420
OY 383 VETIFGSRPMWPGTPRRLPRLPORYMOMRPLFELLGNHACCPGYLLKTCPLRAVT 442
Db 421 PAAGCAREKPOGSAAPAEEDTDPRLVOLLROHSSPMOYGVFACLRRLVPPGLWS 480
OY 443 PAAGCAREKPOGSAAPAEEDTDPRLVOLLROHSSPMOYGVFACLRRLVPPGLWS 502
Db 461 RHNERFLRNTKFTSLGKHAHLSLOELTMKSVDCAMLRSPGVGCYPAEHLREI 540
OY 503 RHNERFLRNTKFTSLGKHAHLSLOELTMKSVDCAMLRSPGVGCYPAEHLREI 562
Db 541 LAKFLHMLMSYVVELLSFEFVYETTFOKNRLFYRKSVMSKLSIGIRHOLKRVOLRE 600
OY 563 LAKFLHMLMSYVVELLSFEFVYETTFOKNRLFYRKSVMSKLSIGIRHOLKRVOLRE 622
Db 601 LSEAEVROHREARPALTSRLRFIPKPDGLRPIVMNDYVVGARTERREKRAERLTSRYKA 660
OY 623 LSEAEVROHREARPALTSRLRFIPKPDGLRPIVMNDYVVGARTERREKRAERLTSRYKA 682
Db 661 LFSVYNTERRARPPGLLGSVGLDIDIRHAMRTFVLRYAODPPPELYFYKVDYGAADI 720
OY 683 LFSVYNTERRARPPGLLGSVGLDIDIRHAMRTFVLRYAODPPPELYFYKVDYGAADI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVOKAAGHVRKAFAKSHVSTLTDLOPYMRQFVAHL 780
OY 743 PODRLTEVIASIIKPONTYCVRRYAVVOKAAGHVRKAFAKSHVSTLTDLOPYMRQFVAHL 802
Db 761 OETSPLRDAVYIEGSSSLNEASSGLFDVFLRFMCHHAVIRKGSYVOCGIPQGSILSTL 840
OY 803 OETSPLRDAVYIEGSSSLNEASSGLFDVFLRFMCHHAVIRKGSYVOCGIPQGSILSTL 862
Db 841 LCSLCYGMENKLFAGIRBDGLLRVLVDDELVTBPHLTHAKFTLRTLVAGVEYICVYNL 900
OY 863 LCSLCYGMENKLFAGIRBDGLLRVLVDDELVTBPHLTHAKFTLRTLVAGVEYICVYNL 922
Db 901 RKTIVNPFVEDEALGTAFAVOMPAHGLFPWCGLLDTRTLEQSDYSYAFRSIASLTF 960
OY 923 RKTIVNPFVEDEALGTAFAVOMPAHGLFPWCGLLDTRTLEQSDYSYAFRSIASLTF 982
Db 961 NRGFRAGNRMRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRPHACVLOLP 1020
OY 983 NRGFRAGNRMRKLFGLVRLKCHSLFLDLQVNSLQTVCTNIYKILLQAYRPHACVLOLP 1042
Db 1021 FHOQWKNPTEFLRYISDTASLCYSLAKKNGMSLGAKGAGPLPSEAVVOMLCHQAFLL 1080
OY 1043 FHOQWKNPTEFLRYISDTASLCYSLAKKNGMSLGAKGAGPLPSEAVVOMLCHQAFLL 1102
Db 1081 KLTRRRTVYVPLGLSLRTAQOTLSRKLPGCTTTALEAANPALPSDFKTIID 1132
OY 1103 KLTRRRTVYVPLGLSLRTAQOTLSRKLPGCTTTALEAANPALPSDFKTIID 1154

```

```

RA MEYERSON M., COUNTER C.M., EATON E.N., ELLISEN L.W., STEINER P.,
RA CADDE S.D., ZIUGRA L., BEIJERSBERGEN R.L., DAVIDOFF M.J., LIU Q.,
RA BACHETTI S., HABER D.A., WEINBERG R.A.,
RL CELL 90:785-795(1997),
DR EMBL; AF018167; G2347129;
SQ SEQUENCE 1132 AA; 126937 MW; C1E5E2AF CRC32;

Query Match      98.1%; Score 8459; DB 4; Length 1132;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 1 MPAPRCRAVNSLSIRSHREVLPATFVRRLGPOGWRVLVYRGDPAARALVAOCLVCPW 60
OY 23 MPAPRCRAVNSLSIRSHREVLPATFVRRLGPOGWRVLVYRGDPAARALVAOCLVCPW 82
Db 61 DARPPAPSTROYSCLELVARVLOLRCERGAKNVLAFGALLDGAARGPPEAFTTSVR 120
OY 83 DARPPAPSTROYSCLELVARVLOLRCERGAKNVLAFGALLDGAARGPPEAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAGMLLRVYDDVLYHLARCALFVLVAPSCAYOYCGPPLVOLA 180
OY 143 SYLPNTVTDALRGSGAGMLLRVYDDVLYHLARCALFVLVAPSCAYOYCGPPLVOLA 202
Db 181 ATOARPPHASGPRRLGECERAMNHSYREAGVPLGLPAPGARRRGGSASRSLPLKRRPR 240
OY 203 ATOARPPHASGPRRLGECERAMNHSYREAGVPLGLPAPGARRRGGSASRSLPLKRRPR 262
Db 241 GAAPERPTPYGQSWAHPRGTRGSPDSGFCVVSAPARPAEATSELEGALSGTRSHSPVG 300
OY 263 GAAPERPTPYGQSWAHPRGTRGSPDSGFCVVSAPARPAEATSELEGALSGTRSHSPVG 322
Db 301 ROHAGPSTSRPPRPMPTPCPPYAEKHFHYSSGDKQOLRPSFLSSLRSLTGARL 360
OY 323 ROHAGPSTSRPPRPMPTPCPPYAEKHFHYSSGDKQOLRPSFLSSLRSLTGARL 382
Db 361 VETIFGSRPMWPGTPRRLPRLPORYMOMRPLFELLGNHACCPGYLLKTCPLRAVT 420
OY 383 VETIFGSRPMWPGTPRRLPRLPORYMOMRPLFELLGNHACCPGYLLKTCPLRAVT 442
Db 421 PAAGCAREKPOGSAAPAEEDTDPRLVOLLROHSSPMOYGVFACLRRLVPPGLWS 480
OY 443 PAAGCAREKPOGSAAPAEEDTDPRLVOLLROHSSPMOYGVFACLRRLVPPGLWS 502
Db 461 RHNERFLRNTKFTSLGKHAHLSLOELTMKSVDCAMLRSPGVGCYPAEHLREI 540
OY 503 RHNERFLRNTKFTSLGKHAHLSLOELTMKSVDCAMLRSPGVGCYPAEHLREI 562
Db 541 LAKFLHMLMSYVVELLSFEFVYETTFOKNRLFYRKSVMSKLSIGIRHOLKRVOLRE 600
OY 563 LAKFLHMLMSYVVELLSFEFVYETTFOKNRLFYRKSVMSKLSIGIRHOLKRVOLRE 622
Db 601 LSEAEVROHREARPALTSRLRFIPKPDGLRPIVMNDYVVGARTERREKRAERLTSRYKA 660
OY 623 LSEAEVROHREARPALTSRLRFIPKPDGLRPIVMNDYVVGARTERREKRAERLTSRYKA 682
Db 661 LFSVYNTERRARPPGLLGSVGLDIDIRHAMRTFVLRYAODPPPELYFYKVDYGAADI 720
OY 683 LFSVYNTERRARPPGLLGSVGLDIDIRHAMRTFVLRYAODPPPELYFYKVDYGAADI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVOKAAGHVRKAFAKSHVSTLTDLOPYMRQFVAHL 780
OY 743 PODRLTEVIASIIKPONTYCVRRYAVVOKAAGHVRKAFAKSHVSTLTDLOPYMRQFVAHL 802
Db 761 OETSPLRDAVYIEGSSSLNEASSGLFDVFLRFMCHHAVIRKGSYVOCGIPQGSILSTL 840
OY 803 OETSPLRDAVYIEGSSSLNEASSGLFDVFLRFMCHHAVIRKGSYVOCGIPQGSILSTL 862
Db 841 LCSLCYGMENKLFAGIRBDGLLRVLVDDELVTBPHLTHAKFTLRTLVAGVEYICVYNL 900
OY 863 LCSLCYGMENKLFAGIRBDGLLRVLVDDELVTBPHLTHAKFTLRTLVAGVEYICVYNL 922
Db 901 RKTIVNPFVEDEALGTAFAVOMPAHGLFPWCGLLDTRTLEQSDYSYAFRSIASLTF 960

```



```

QY 923 RTIVNPFVEDEALGTAIVQPAHGLFPWCGLLDTRTLEVOSSDYASRTSIRASLTF 982
Db 961 NNGFKAGMRRRLRGLVLRKCHSLFDLOVNSLOFVCNITYKLLLOAYRPAQLOLP 1020
QY 983 NNGFKAGMRRRLRGLVLRKCHSLFDLOVNSLOFVCNITYKLLLOAYRPAQLOLP 1042
Db 1021 FHOQVKNPFFELRVISDPTASLCYSILKAKNAGMSLGAGAGPLPSEAVQWLCHQAFLL 1080
QY 1043 FHOQVKNPFFELRVISDPTASLCYSILKAKNAGMSLGAGAGPLPSEAVQWLCHQAFLL 1102
Db 1081 KLTRHRVTVPLLSGLRTAQOTLSRRLPGTTLTALBAANPALPSDFKTLTD 1132
QY 1103 KLTRHRVTVPLLSGLRTAQOTLSRRLPGTTLTALBAANPALPSDFKTLTD 1154

RESULT 3
ID 070372 PRELIMINARY; PRT; 1122 AA.
AC 070372;
DT 01-AUG-1998 (TREMBLERL. 07, CREATED)
DT 01-AUG-1998 (TREMBLERL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLERL. 07, LAST ANNOTATION UPDATE)
OS MUS MUSCULUS (MOUSE).
OC TELOMERASE REVERSE TRANSCRIPTASE.
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
RN [1]
RP SEQUENCE FROM N.A.
RA GREENBERG R.A., ALLSOPP R.C., CHIN L., MORIN G.B., DEPINHO R.A.;
RL ONCOGENE 0:0-0(1998).
DR EMBL, AF051911; G3005592; -.
KW RNA-DIRECTED DNA POLYMERASE.
SQ SEQUENCE 1122 AA; 127977 MW; 22207506 CRC32;

Query Match 59.8%; Score 5161; DB 11; Length 1122;
Best Local Similarity 62.7%; Pred. No. 0.00e+00;
Matches 718; Conservative 198; Mismatches 192; Indels 38; Gaps 24;

Db 1 MRRAPRCVAVRSILSRSEYEWPLAFVRLGEGRRLOVQPGDKYIRTYACQICMNM 60
QY 23 MRRAPRCVAVRSILSRSEYEWPLAFVRLGEGRRLOVQPGDKYIRTYACQICMNM 82
Db 61 GSOPPRADLSFHQVSSILKELVAVVOVRLCERNRNVLAFGEFELNBARGGPPMAFTSSVR 120
QY 83 DARPPRAASFRQVSLKELVAVVOVRLCERNRNVLAFGEFELNBARGGPPMAFTSSVR 142
Db 121 STLPNTVIELRVSGAMLLSRVGGDLVYLLAHACALYLVPPSCAYOYCGSPLYQICA 180
QY 143 STLPNTVIELRVSGAMLLSRVGGDLVYLLAHACALYLVPPSCAYOYCGSPLYQICA 202
Db 181 TTDIMPSVASIRPTPVGRNFTNLFLOQIKSSROEAPKPLALPSRGKRLSLTSTS 240
QY 203 ATQARPPRAAS-GPRRLGCE-RA-W--NH--SVR-EGVPLGPGARRRGSGASAS 253
Db 241 VSARKARCVPRVPRVEGPHROVLPTRPSGKSWPSPARS-PEV-PT-AEKDLS-S-KGYS 296
QY 254 LRLPKPRRGARPERPRTRYGOGSMAHPGRTGPS-DRGCVYSPARPPEATSLGALS 312
Db 297 DLSLS-GSYCKHKRPSSTLSLSPRONAQLRP-FLETIRHFLSRDGOERLNPSTLSN 354
QY 313 GTRHSHPSVGRQHNAGRPSTSRPRPMWPCRPVYATKTKFLVSSD-KQLRPSLSS 371
Db 355 LQPNLTGARLYEITFLGSRPTSGPLCRNHSRKRYQMWRPFLFOOLLVNHACQVRL 414
QY 372 LRPSLTGARLYEITFLGSRPTSGPLCRNHSRKRYQMWRPFLFOOLLVNHACQVRL 431
Db 415 RSHCRPR--T--AN-----O--O--VT--DALUTSPHLMDLRLSSPWQYVGRACI 459
QY 432 KTHCPRLAAVTPRPAVCANRKPQGSVAAPREEDIDRRRLVQLRKHSSPWQYVGRACI 491
Db 460 CKVVSASLNGTRNRRRFFKRLKFLSLGKYGLSLQELMMKKMKVEDCHWLRSPPKDRY 519
QY 492 RRLVPPGLMGSRRNRRRLRNTKTKFLSLGKHAKLSLQELMMKKMKVEDCHWLRSPPKDRY 551

```

```

Db 520 PAEHRRLREILATFELWIMDYVOLLRSFFYITESTPOKNRLEFFRYKSVMSKQISGV 579
QY 552 PAEHRRLREILAKFLHWMISYVVELLSFFYIETTPQKRRLEFFRYKSVMSKQISGI 611
Db 580 ROHLERVLRELSOEVRNHOOTLAMPICRLPFLPKPGLRPYNNMSYSGTALGRK 639
QY 612 ROHLKRVQLRELSAEVRNHOOTLAMPICRLPFLPKPGLRPYNNMSYSGTALGRK 671
Db 640 QAOHFORLKTFLSMINTERTKHPHLMGSSVYGMNDIYRTMAFYLRYVALDOTPRMYTV 699
QY 672 RAERLTSRKALFESVLYNTERARRPGLLGASVYGLDIHRAMTFYLVRYAODPPRELYV 731
Db 700 KADVTGAYATPOGKLYEVVANNIRSESTYCIROYAVVRROGCVHKSFRROYTLSD 759
QY 732 KYDTYCATYITPODRLEVIASLIK-PQNTYCYKRAVYQKAAHGVKRAFSHSTLTD 790
Db 760 LQPYMGQFLKHLQDSDASALRNSVYIEQISNMSSSSLFDEFLEHRLSHSVYKIGRCYT 819
QY 791 LQPYMGQFLKHLQDSDASALRNSVYIEQISNMSSSSLFDEFLEHRLSHSVYKIGRCYT 848
Db 820 QCOGIPQGSSTLTLCSLCFCGEMENKLPFAVQRODGLLRFVDEFLVTHLDOAKTFLST 879
QY 849 QCOGIPQGSSTLTLCSLCFCGEMENKLPFAVQRODGLLRFVDEFLVTHLDOAKTFLST 908
Db 880 LVHGVPEYGCMTLQKTYVNFPEVPEPTLAGAAPYOLPAHCPFLPMCGLLDPTQLEVFCDY 939
QY 909 LVHGVPEYGCMTLQKTYVNFPEVPEPTLAGAAPYOLPAHCPFLPMCGLLDPTQLEVFCDY 968
Db 940 SGYATQISITSLTFOSVFRAGTMRKLLSVLRCHGLFDLOVNSLOFVCNITYKIEL 999
QY 969 SGYATQISITSLTFOSVFRAGTMRKLLSVLRCHGLFDLOVNSLOFVCNITYKIEL 1028
Db 1000 LQAYRFAVYQIPEQDQARKULTFPLGLIISQASQCIYILKVPKNGML-K-ASGSP 1056
QY 1029 LQAYRFAVYQIPEQDQARKULTFPLGLIISQASQCIYILKVPKNGML-K-ASGSP 1088
Db 1057 PEAAHMLCYQAFELKLAASVYLYCKLGLPRTAOKLCKRLPEATWTILKAADPALSD 1116
QY 1089 PEAAHMLCYQAFELKLAASVYLYCKLGLPRTAOKLCKRLPEATWTILKAADPALSD 1148
Db 1117 FQTIID 1122
QY 1149 FQTIID 1154

RESULT 4
ID 013339 PRELIMINARY; PRT; 988 AA.
AC 013339;
DT 01-JAN-1998 (TREMBLERL. 05, CREATED)
DT 01-JAN-1998 (TREMBLERL. 05, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLERL. 07, LAST ANNOTATION UPDATE)
OS TELOMERASE REVERSE TRANSCRIPTASE 1.
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCMYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RA LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;
RL MEDLINE; 97400623.
DR NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
QY LINNEN J., HARLEY C.B., CECI T.R.;
QY SCIENCE 277:955-959(1997).
QY [2]
QY SEQUENCE FROM N.A.
QY STRAIN-972H-;
QY LYNE M., RAJANDREAM M.A., BARRELL B.G., VOLCKAERT G.;
QY SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
QY EMBL; AF015783; G2340168; -.
QY EMBL; AL022299; E1285360; -.
QY PFAM; PF00078; tvt.
QY RNA-DIRECTED DNA POLYMERASE.
QY SEQUENCE 988 AA; 116328 MW; 39C385A7 CRC32;

```

QY	540	AML--RSPGVGCPAHEHRLREELIAKFLHMLMSYVVELLSFFVYETTFQKKRLF	596
Db	458	YFRKDIW-KLLCRPFTSMKMEAEKINENNVRMDTO-KTLLPFAVIRLLPKKNTFPLIT	515
QY	597	YFRKSWSKLOSIGIRHQLKRVQLRELSAEVROHREARPALITSRLRFTPKPGDLPIV	656
Db	516	NLRKFFLLKQGSNK-KMLVSTNQTLRPVLSILHNLINSSSGIPFN-LEVMMLLMEFK	573
QY	657	NM-D-YVVGANTFRERKEAERLTSRVKALFSEVLNTERARRGGLLGASVLGDDIHRIWRT	714
Db	574	DLKRRMFGRK-K-YEVAIDIKSCYDRIKODLMEIRYAKKLKDE-EVIRKATIH-ATS	629
QY	715	FVLRADDPPELLEFYKVDYTGAVDPIRQDLREVLASIKPKONTQCVRRYAVVQGAH	774
Db	630	DRATKFEVSEAFYFDWPFKE--VOLLMSKT-SDTLFVDYVDWTKSSSEIFKMLKEH	686
QY	775	GHVRAKFSHSVSTLTDLPYKRFVYAHQETSPLRDVAIVEOSSLSLNEASSGLDFELRF	834
Db	667	LSGHTVTKGNSQYLOKVGILPGGSLSSLCHEFYMEDLIDELSTKKGKGSYLLRVYDFL	746
QY	835	MCHAVNRIRGKSYVOCQIGIPGSLSTLCSLCYDMEKMLFAGIRBDG-LLLRLVDFL	893
Db	747	FTYVKKKAKKFLNLSLGEFKHNFSTLEKTVFNFSNGIINNTFENESKR-M-PFF	804
QY	894	LVTPLHTAKFTFLRLVKGVPDYGCYVNLKRTVYVNFVEDALGCTAFVQMPAHGLPWC	953
Db	805	GFSVNMRSLD 814	
QY	954	GILDTFRLLE 963	
RESULT	6		
ID	000939	PRELIMINARY;	PRT; 1031 AA.
AC	000939		
DT	01-JUL-1997	(TREMBLREL. 04, CREATED)	
DT	01-JUL-1997	(TREMBLREL. 04, LAST SEQUENCE UPDATE)	
DT	01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)	
DE	TELOMERASE SUBUNIT P123.		
OS	EUPLOTES AEDICULATUS.		
OC	EUKARYOTAL; MITOCHONDRIAL; EUKARYOTES; ALVEOLATA; CILIOPHORA;		
OC	HYPOTRICHs; EUPLOTIDA; EUPLOTES.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE; 97274210.		
RA	LINGNER T., HUGHES T.R., SHEVCHENKO A., MANN M., LUNDBLAD V.,		
RA	CECH T.R.,		
RL	SCIENCE 276:561-567(1997).		
DR	EMBL; U95964; G2072336.		
SO	SEQUENCE 1031 AA; 122562 MW; 214885CD CRC32;		
Query Match	5.1%;	Score 439;	DB 5; Length 1031;
Best Local Similarity	23.0%;	Pred. NO. 1.97e-60;	
Matches 14;	Conservative 168;	Mismatches 261;	Indels 47; Gaps 34
Db	361	FINEFYNILPRDELITGR-NRKNFQKKYKTVLKNHE-LIHKMLLEKINTREISMVQV	418
QY	486	FVRACTRLRLVPPGLMGSRHNRREFLRNTKFLISLGHAKLSLQELTW-KMSVDCAMLLR	544
Db	419	EISAHFFYFDHE-NIYVLKMLNWFIEDLVVSLIRCFYTTBEOOKSYSKYVRRKNTMD	477
QY	545	SPGVGCPAHEHRLREELIAKFLHMLMSYVVELLSFFVYETTFQKKRLFYKRSWS	604
Db	478	VIMKSI-ADLKEKTELAEVQKEV-EWKRSLGFAFGKRLIPKTFRPI--MTF--NK	531
QY	605	KLOSIGIRHQLKRVQLRELSAEVROHREARPALITSRLRFTPKPGDLPIVNMDDYVGA	664
Db	532	KIVNSDRRTKTLTNTKLLNSHMLKTLKNMFDPGFAVNTDDVWKYEEFVCK-WK	590
QY	665	RFRERKEAERLTSRVKALF--VLNTERARR-DGLGASVGLDDIHRAARTFVLRYRA	721
Db	591	QVGGKLPFAIMDIEKCYDSVNRKLSFTLTLLSSDFEIMTAQILKRNNTVIDSKN	650
QY	722	ODPPELLEFYVADYTGAVDPIRQDLREVLASIKPKONTQCVRRYAVVQGAH-Q-KA	772

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 01, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN (FRAGMENT).
OS HUMAN HERPESVIRUS TYPE 4.
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE;
OC GAMMAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87284169.
RA PRITZNER A.J., TSAI E.C., STROMINGER J.L., SPECK S.H.;
RL J. VIROL. 61:2902-2909(1987).
DR EMBL; M17294; G807646; -.
KM HYPOTHETICAL PROTEIN.
FT NON_TER 1
SQ SEQUENCE 296 AA; 31393 MW; D17CE6F CRC32;

Query Match 1.7%; Score 147; DB 14; Length 296;
Best Local Similarity 38.6%; Pred. No. 7,18e-06;
Matches 39; Conservative 20; Mismatches 33; Indels 9; Gaps 8;

Db 91 GPPPTSGAAQ-RTNRPPGCPRSANPCCPTWRR--R-SGAQKHPPPGAGQPSGP 146
QY 194 GPPYQLGAATQARPPhASG-PRR-R-IGCERAMNHSVREAGVPLGAPGA-RRRGS 249
Db 147 TGGRRAPGAPGTAPAGPGGGAIVPSGATPHRRSGSPAD 187
QY 250 ASRSLPLKRRPRRGAPEPER-TEVGQGSMAHPGRTGSPD 289

RESULT 10
ID 036027; PRELIMINARY; PRT; 574 AA.
AC 036027;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE WISKOT-ALDRICH SYNDROME PROTEIN HOMOLOG 1.
GN WSP1 OR SPAC4F10.15C
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JS21;
RA ZANKEL T.C., OW D.W.;
RL SUBMITTED (DEC-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN-972;
RA CONNOR R., CHURCHER C.M., BARRELL B.G., RAJANDREAM M.A., WOOD V.;
RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- SIMILARITY: TO YEAST LAS17.
DR EMBL; AF038575; G2708709; -.
FT DOMAIN 311 317 POLY-PRO.
FT DOMAIN 337 343 POLY-PRO.
FT DOMAIN 361 366 POLY-PRO.
FT VARIANT 248 248 L -> V (IN STRAIN JS21).
SQ SEQUENCE 574 AA; 59605 MW; 9E84D37C CRC32;

Query Match 1.7%; Score 143; DB 3; Length 574;
Best Local Similarity 28.8%; Pred. No. 2,85e-05;
Matches 42; Conservative 41; Mismatches 53; Indels 10; Gaps 7;

Db 305 AANKRRPPPP-PSRR--NRG-KPPIGNSSNSLPPPPPSNAAG-SIPLPQGR 358
QY 202 AATGAPRRPHASGRRLRGCRANHSVREAGVPLGAPGARRRGSGASLPLPRR 261
Db 359 SAPPPPRSAPSTGKRPPLSSRAVSNP--APPALPKRSAPALPLGNASTSTPP 416
QY 262 RGAAPDEPRTPVGGSWAHP-GRTRGPSDRGFCVSPARPAEATSLGALSTGRSHSPS 320
Db 417 VPTPP-SLPPSAPSLPPSAPSLPM 441

QY 321 VGRQHNAGPSTSRPRPMDTCCPV 346

RESULT 11
ID 000600; PRELIMINARY; PRT; 234 AA.
AC 000600;
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JUL-1997 (TREMBLREL. 04, LAST ANNOTATION UPDATE)
DE PAROTID 'O' PROTEIN (FRAGMENT).
GN PRB4.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96108975.
RA AZEN E.A., AMBERGER E., FISHER S., PRAKOBPHOL A., NIEGE R.L.;
RL AM. J. HUM. GENET. 58:143-153(1996).
DR EMBL; S80916; G1911492; -.
FT NON_TER 1
SQ SEQUENCE 234 AA; 23656 MW; B5920075 CRC32;

Query Match 1.6%; Score 142; DB 4; Length 234;
Best Local Similarity 28.2%; Pred. No. 4,02e-05;
Matches 40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;

Db 92 QGQNSQG-PPPHGRPRPP-OGG-NOSHRPPPP-GKPER-PPPGQNSQG-P-PP 144
QY 199 QLGATQARPPPHASGPRRLGCRAMNHSVREAGVPLGAPGARRRGSGASLPLPK 258
Db 145 HPGKPEPPPOEKNKSASPPGKQPGQDGNPQPGPPGKQPGPPGNGNQOQ 204
QY 259 RPRGAPEPERTPVGGSWAHPGRTGPSR-GFCVSPAPAE-EAUSLEGALSGLTR 316
Db 205 APPAGKPGPPPPGGRPPR 226
QY 317 SHPSVGRQHNAGPSTSRPPR 338

RESULT 12
ID 046612; PRELIMINARY; PRT; 276 AA.
AC 046612;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE IS 1222 GENE ORF-A AND ORF-B.
OS ENTEROBACTER AGGLOMERANS.
OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS;
OC ENTEROBACTERIACEAE.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-KLEBERGER, 1983;
RX MEDLINE; 9525664.
RA STEIBL H.D., LEMECKE F.M.;
RL GENE 156:37-42(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA STEIBL H.D., SIDDAVATTAM D.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 96422758.
RA STEIBL H.D., SIDDAVATTAM D., KLINGMUELLER W.;
RL PLASMID 34:223-228(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA STEIBL H.D.;
RL THESES (1995), UNIVERSITAET BAYREUTH.
DR EMBL; X78052; G459248; -.
DR EMBL; X81893; E258949; -.
SQ SEQUENCE 276 AA; 31718 MW; 0B54A420 CRC32;

```

Query Match          1.6%:   Score 141;   DB 2; Length 276;
Best Local Similarity 33.7%:   Pred.No. 5,65e-05;
Matches      30; Conservative    23; Mismatches 30; Indels    6; Gaps    6;

Db       42 ITTELERR-FGYRR-TMOLREGILVHNKRYRRLTHSLGVKKRRR-RKGLATERL 98
|||:::||||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy       585 VTETFQNNRLFYFKXSWSKLSGIGNHLKRV-QLEHSEAEAYRQHREARNALLISRL 643
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db       99 PLL-RPAAPNLTWSDPEVMALALGRRIK 126
:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Oy       644 REIPKPDGLRPVNMDYVGA-RIFRRERK 671

RESULT      13
ID           099076        PRELIMINARY;      PRF;      316 AA.
AC           099076;

DT         01-NOV-1996 (TREMBLEREL. 01, CREATED)
DT         01-NOV-1996 (TREMBLEREL. 01, LAST SEQUENCE UPDATE)
DT         01-JUN-1998 (TREMBLEREL. 06, LAST ANNOTATION UPDATE)
DE         HOMEBOX PROTEIN (FRAGMENT).
GN         HB9.
OS         HOMO SAPIENS (HUMAN).
OC         EUCHAROTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC         EUTHERIA; PRIMATES.
RN         [1]
RP         SEQUENCE FROM N.A.
RC         TISSUE-TONSIL;
RX         MEDLINE; 91305125.
RA         DEGUCHI Y., KEHR J.H.;
RL         NUCLEIC ACIDS RES. 19:3742-3742(1991).
CC         -I- SIMILARITY: STRONG WITH OTHER "ANTP-TYPE" HOMEBOX PROTEINS.
DR         EMBL; X56537.E30256; ALT_SEQ.
PR         PF00046; S16681;
RW         DNM-BINDING; NUCLEAR PROTEIN; HOMEBOX; PHOSPHORYLATION.
FT         DNM_BIND      1      NON_TER
FT         DNA_BIND     70     129     HOMEBOX.
FT         DN_BIND     112     121     H-T-H MOTIF.
SQ         SEQUENCE      316 AA; 34713 MW; 1531E766 CRC32;

Query Match          1.6%:   Score 136;   DB 4; Length 316;
Best Local Similarity 30.9%:   Pred.No. 3,03e-04;
Matches      21; Conservative    20; Mismatches 24; Indels    3; Gaps    3;

Db       202 RFWTASDSCVGRTGIRPGPAT-SPPSSSPAQDSRARBARPAPSPGSGAWTH 260
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Oy       223 RAWNVSTVEAGV-PLGLPAPGARRGGSASRLPLPKRPARG-AAPEPTPVGGSMWH 280
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Db       261 PARPREA 268
|||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

Oy       281 PORTGPS 288

RESULT      14
ID           054226        PRELIMINARY;      PRF;      373 AA.
AC           054226;

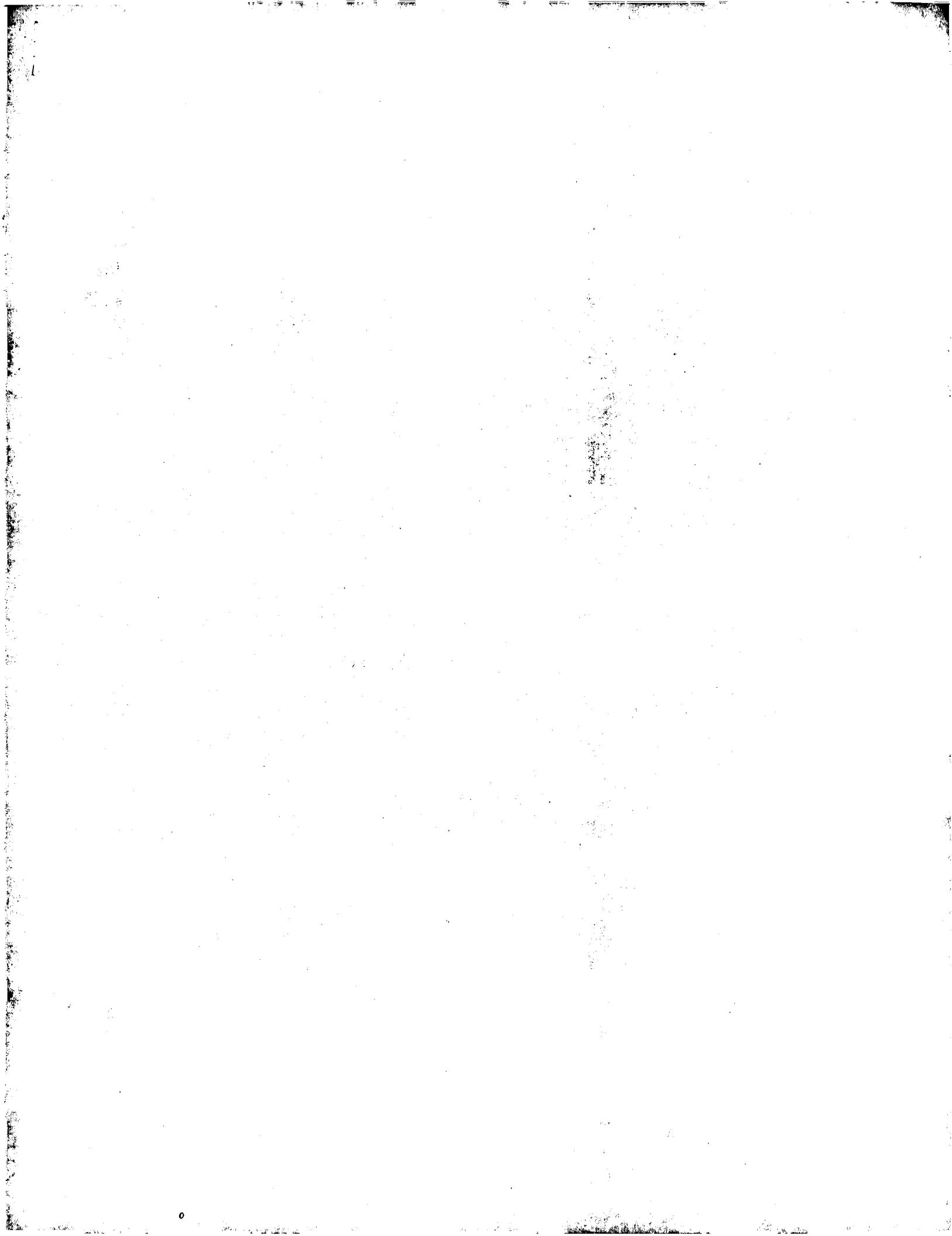
DT         01-JUN-1998 (TREMBLEREL. 06, CREATED)
DT         01-JUN-1998 (TREMBLEREL. 06, LAST SEQUENCE UPDATE)
DT         01-AUG-1998 (TREMBLEREL. 07, LAST ANNOTATION UPDATE)
DE         POLYKETIDE SYNTHASE (FRAGMENT).
GN         ERYA.
OS         SACCHAROPOLYSORA ERYTHRAEA (STREPTOMYCES ERYTHRAEUS).
OC         PROKARYOTA; FIRMICUTES; ACTINOMYCETALES; NOCARDIOFORM.
RN         [1]
RP         SEQUENCE FROM N.A.
RC         STRAIN-NRRL2338;
RA         SALAH-BEY K., DOUMITH M., MICHEL J.M., HAYDOCK S., CORTES J.,
LEADLAY P.F., RAYNAL M.C.;
RL         SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR         EMBL; Y14332; E1228185; -.
FT         NON_TER      1      1
SQ         SEQUENCE      373 AA; 40872 MW; 781BC6D CRC32;
```

	Query Match	1.6%	Score 140;	DB 2;	Length 373;
	Best Local Similarity	32.38%	Pred. No. 7.93e-05;		
	Matches	40;	Conservative	27;	Mismatches 46; Indels 11; Gaps 11.
Db	243 PRRAA-NRRAG-AAPAPGIAPG-AAHG-HRRG-AARRARDRRRGGGRGRRGOPRPG	297	: : : : : : : : : : :		
QY	215 PRRRCGRANNNHVSREKGVPLGLPAPCARRRGSASLSULPRPRRGAAPEPERTPVG	274	: :		
Dd	236 L-RRARPPR-RGPPIRRRAVGTSHRPPOACG-AGH-R-ARHRTA-GRGOGAARTTAQR	351	: : : : : : : : : : : : : : : : : : :		
QY	275 QGSMVAHPERTGRPSDRGCVCVPSPARPAPEAATSLGALSTGRHSHP SVGRHHADPSTSR	334	: :		
Dd	352 PGRP 355				
QY	335 PPPP 338				

	RESULT	15	PRELIMINARY;	PRI;	585	AA.
AC	ID	041935				
AC		041935.				
DT		01-JAN-1998	(TREMBLREL. 05, CREATED)			
DT		01-JAN-1998	(TREMBLREL. 05, LAST SEQUENCE UPDATE)			
DT		01-JAN-1998	(TREMBLREL. 05, LAST ANNOTATION UPDATE)			
DE		HYPOTHETICAL 60.2 KD PROTEIN.				
CN		GAMMAHV.M6.				
OS		MORVINE HERPESVIRUS 68.				
OC		VIRUSES; DSDNA VIRUSES, NO RNA STAGE; HERPESVIRIDAE; GAMMAHERPESVIRINAE				
RN		[1]				
RP		SEQUENCE FROM N.A.				
RC		STRAIN-WDMS;				
RC		MEDLINE; 97366649.				
RA		VIRGIN H.W. IV, LATREILLE P., WAMSLEY P., HALLSWORTH K., WECK K.E.,				
RA		DAL CANTO A.J., SPECK S.H.;				
RL		J. VIROL. 71:5894-5904(1997).				
RN		[2]				
RP		SEQUENCE FROM N.A.				
RC		STRAIN-WDMS;				
RC		LATREILLE P., WAMSLEY P., WATERSTON R.H.;				
RL		SUBMITTED (APR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.				
DR		EMBL: U97553; G3317934; -				
KW		HYPOTHETICAL PROTEIN.				
SO		SEQUENCE	585	AA;	60160	MM; B35C72E5 CRC32;

	Query Match	1.6%	Score 135	DB 14	Length 585
	Best Local Similarity	27.6%	Pred. No. 4.22e-04		
	Matches	45	Conservative	37	Mismatches 68; Indels 13; Gaps 11
Db	342 PPP-ELGPGSTSPASRAGARIPDLGPIPLSWCPDPRPPRPPELGSPGSPSPASRAG	400			
Oy	195 PPIYOLGAATQARRPPHNASGPRRR-L-GCEAHMHSVREA-GYV-LDLPPAGARRRGSA	250			
Db	401 ARIPDLGPIPLSWCPDPRPPRPPELG-GSPY-SPASRAGARIIPRSPRPPELG-G	456			
Oy	251 SRSLPLPARR-PRRGAAPERTPTVGOGSGMARRGTRGSPDSGFCVSPARAEATSLG	309			
Db	457 SPRPRPSRRP-PELGSGPDLGPIPLSWG-PRRPFPAASR	496			
Oy	310 ALSGTRHSHPSVGRONHAGRPSTRRPAPWTCRPPVAAETKH	352			

Search completed: Fri Dec 18 18:30:12 1998
Job time : 190 secs.



WORLDWIDE
(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:30:31 1998; MasPar time 9.93 Seconds
Tabular output not generated. 823.071 Million cell updates/sec

Title: >US-08-951-733-20
Description: (1-1154) from US08951733.pep
Perfect Score: 8624
Sequence: 1 HASGQRCVLRTWEALAPAT.....TALEAANPALPSDFETILD 1154

Scoring table: PAM 150
Gap 11

Searched: 77309 seqs, 7078906 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:S_COMB 2:PCT9_COMB 3:backfilest1

Statistics: Mean 37.265; Variance 182.038; scale 0.205

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	118	1.4	399	1	US-08-147-Sequence 27, Applicati	2.06e+00
2	118	1.4	399	1	US-08-447-Sequence 27, Applicati	2.06e+00
3	118	1.4	399	2	PCT-US92-0-Sequence 23, Applicati	2.06e+00
4	118	1.4	399	1	US-08-278-Sequence 23, Applicati	2.06e+00
5	118	1.4	399	1	US-08-643-Sequence 23, Applicati	2.06e+00
6	118	1.4	399	1	US-07-901-Sequence 13, Applicati	2.06e+00
7	118	1.4	399	1	US-08-643-Sequence 23, Applicati	2.06e+00
8	118	1.4	399	1	US-08-479-Sequence 10, Applicati	2.06e+00
9	118	1.4	399	1	US-08-155-Sequence 23, Applicati	2.06e+00
10	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
11	118	1.4	399	2	PCT-US93-1-Sequence 13, Applicati	2.06e+00
12	118	1.4	399	2	PCT-US93-0-Sequence 10, Applicati	2.06e+00
13	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
14	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
15	118	1.4	399	2	PCT-US93-0-Sequence 8, Applicati	2.06e+00
16	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
17	118	1.4	399	1	US-08-480-Sequence 10, Applicati	2.06e+00
18	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
19	118	1.4	399	1	US-08-406-Sequence 23, Applicati	2.06e+00
20	118	1.4	399	1	US-08-451-Sequence 23, Applicati	2.06e+00
21	118	1.4	399	2	PCT-US93-0-Sequence 23, Applicati	2.06e+00
22	118	1.4	399	2	PCT-US91-0-Sequence 6, Applicati	2.06e+00
23	125	1.4	402	1	US-07-841-Sequence 29, Applicati	7.05e-01

24	125	1.4	402	1	US-08-643-Sequence 21, Applicati	7.05e-01
25	125	1.4 <th>402</th> <th>1</th> <th>US-08-451-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-451-Sequence 21, Applicati	7.05e-01
26	125	1.4 <th>402</th> <th>1</th> <th>US-08-462-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-462-Sequence 21, Applicati	7.05e-01
27	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 21, Applicati	7.05e-01
28	125	1.4 <th>402</th> <th>1</th> <th>US-08-206-Sequence 4, Applicati</th> <th>7.05e-01</th>	402	1	US-08-206-Sequence 4, Applicati	7.05e-01
29	125	1.4 <th>402</th> <th>1</th> <th>US-08-155-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-155-Sequence 21, Applicati	7.05e-01
30	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 21, Applicati	7.05e-01
31	125	1.4 <th>402</th> <th>1</th> <th>US-08-480-Sequence 8, Applicati</th> <th>7.05e-01</th>	402	1	US-08-480-Sequence 8, Applicati	7.05e-01
32	125	1.4 <th>402</th> <th>1</th> <th>US-07-901-Sequence 11, Applicati</th> <th>7.05e-01</th>	402	1	US-07-901-Sequence 11, Applicati	7.05e-01
33	125	1.4 <th>402</th> <th>1</th> <th>US-08-147-Sequence 29, Applicati</th> <th>7.05e-01</th>	402	1	US-08-147-Sequence 29, Applicati	7.05e-01
34	125	1.4 <th>402</th> <th>1</th> <th>US-08-643-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-643-Sequence 21, Applicati	7.05e-01
35	125	1.4 <th>402</th> <th>1</th> <th>US-08-406-Sequence 8, Applicati</th> <th>7.05e-01</th>	402	1	US-08-406-Sequence 8, Applicati	7.05e-01
36	125	1.4 <th>402</th> <th>1</th> <th>US-08-479-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-479-Sequence 21, Applicati	7.05e-01
37	125	1.4 <th>402</th> <th>1</th> <th>US-08-278-Sequence 29, Applicati</th> <th>7.05e-01</th>	402	1	US-08-278-Sequence 29, Applicati	7.05e-01
38	125	1.4 <th>402</th> <th>1</th> <th>US-08-447-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	1	US-08-447-Sequence 21, Applicati	7.05e-01
39	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 21, Applicati	7.05e-01
40	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 21, Applicati	7.05e-01
41	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-1-Sequence 8, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-1-Sequence 8, Applicati	7.05e-01
42	125	1.4 <th>402</th> <th>2</th> <th>PCT-US92-0-Sequence 21, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US92-0-Sequence 21, Applicati	7.05e-01
43	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 11, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 11, Applicati	7.05e-01
44	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 6, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 6, Applicati	7.05e-01
45	125	1.4 <th>402</th> <th>2</th> <th>PCT-US93-0-Sequence 6, Applicati</th> <th>7.05e-01</th>	402	2	PCT-US93-0-Sequence 6, Applicati	7.05e-01

ALIGNMENTS

RESULT 1
ID US-08-147-023-27 STANDARD; PRT; 399 AA.
XX
AC
XX
xx
xx

Sequence 27, Application US/08147023

Sequence 27, Application US/08147023

Patent No. 5468845

GENERAL INFORMATION:

APPLICANT: OPPERMAN, HERMANN

APPLICANT: OZKAYNAK, ENGIN

APPLICANT: KUBERASAMPATH, THANGAVEL

APPLICANT: RUEGER, DAVID C.

APPLICANT: PANG, ROY H.L.

TITLE OF INVENTION: OSTEOGENIC DEVICES

NUMBER OF SEQUENCES: 33

CORRESPONDENCE ADDRESS:

ADDRESSEE: TESTA, HURWITZ & THIBEAULT

STREET: 53 STATE STREET

CITY: BOSTON

STATE: MASSACHUSETTS

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/147,023

FILING DATE: 21-FEB-1992

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 810,560

FILING DATE: 20-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 827,052

FILING DATE: 28-JAN-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 660,162

FILING DATE: 22-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 621,988

FILING DATE: 04-DEC-1990

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 621,849
CC FILING DATE: 04-DEC-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 616,374
CC FILING DATE: 21-NOV-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 600,024
CC FILING DATE: 18-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 599,543
CC FILING DATE: 18-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 579,865
CC FILING DATE: 07-SEP-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 569,920
CC FILING DATE: 20-AUG-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 483,913
CC FILING DATE: 22-FEB-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 422,613
CC FILING DATE: 17-OCT-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 315,342
CC FILING DATE: 23-FEB-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 232,630
CC FILING DATE: 15-AUG-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 179,460
CC FILING DATE: 08-APR-1988
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER, EDMUND R.
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-001CP6
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/248-7000
CC TELEFAX: 617/248-7100
CC INFORMATION FOR SEQ ID NO: 27:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;
SQ
Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Fred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;
Db 12 GLACALGGHGRPP-HTC-PORRLG-ARERDMOREILLAVLGLPGR-PPRAQPAAR 67
Qy 194 GPPLXQGAATQAPPPHASPRLRGCEKAMNSVREAGVPLGLPAGARRGGSASRS 253
Db 68 QP 69
Qy 254 LP 255

CC APPLICANT: OPPERMAN, HERMAN
CC APPLICANT: OKAYAK, ENGIN
CC APPLICANT: KOBERSAPATH, THANGAVEL
CC APPLICANT: ROGER, DAVID C.
CC APPLICANT: PANG, ROY H. L.
CC TITLE OF INVENTION: OSTEOGENIC DEVICES
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: TESTA, HORWITZ & THIBEAULT
CC STREET: 53 STATE STREET
CC CITY: BOSTON
CC STATE: MASSACHUSETTS
CC COUNTRY: U.S.A.
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/447,570
CC FILING DATE: 21-FEB-1992
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 810,560
CC FILING DATE: 20-DEC-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 827,052
CC FILING DATE: 28-JAN-1992
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 660,162
CC FILING DATE: 22-FEB-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 621,988
CC FILING DATE: 04-DEC-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 621,849
CC FILING DATE: 04-DEC-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 616,374
CC FILING DATE: 21-NOV-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 600,024
CC FILING DATE: 18-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 599,543
CC FILING DATE: 18-OCT-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 579,865
CC FILING DATE: 07-SEP-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 569,920
CC FILING DATE: 20-AUG-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 483,913
CC FILING DATE: 22-FEB-1990
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 422,613
CC FILING DATE: 17-OCT-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 315,342
CC FILING DATE: 23-FEB-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 232,630
CC FILING DATE: 15-AUG-1988
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 179,460
CC FILING DATE: 08-APR-1988
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER, EDMUND R.
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-001CP6
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 617/248-7000
CC TELEFAX: 617/248-7100
CC INFORMATION FOR SEQ ID NO: 27:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA: 44764 MW: 790568 CN;

Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GIALCALGGHGPRPP-HTC-PQRLG-ARERRDMOREILAVLPGR-PPRPAQPAAR 67
QY 194 GPPLYOLGATQARPPPHASGPRRLGCRERAMNHSVREAGVPLGLPAPGARRRGSGSASRS 253

Db 68 QP 69
QY 254 LP 255

RESULT 3
ID PCT-US92-01968-23 STANDARD; PRT; 399 AA.
XX xxxxxx
DT
XX
DE Sequence 23, Application PC/TUS9201968

CC Sequence 23, Application PC/TUS9201968
CC GENERAL INFORMATION:
CC APPLICANT: COHEN, CHARLES M
CC APPLICANT: KUBERASAMPATH, THANGAVEL
CC APPLICANT: PANG, ROY H L
CC APPLICANT: OPPERMANN, HERMANN
CC APPLICANT: RUEGER, DAVID C
CC TITLE OF INVENTION: PROTEIN-INDUCED MORPHOGENESIS
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: TESTA HURWITZ & THIBEAULT
CC STREET: EXCHANGE PLACE 53 STATE STREET
CC CITY: BOSTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02109
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US92/01968
CC FILING DATE: 19920311
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 667,274
CC FILING DATE: 11-MAR-1991
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 752,764
CC FILING DATE: 30-AUG-1991
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER ESO, EDMUND R
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-052PC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617/248-7000
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: AMINO ACID

CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA: 44764 MW: 790568 CN;

Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GIALCALGGHGPRPP-HTC-PQRLG-ARERRDMOREILAVLPGR-PPRPAQPAAR 67
QY 194 GPPLYOLGATQARPPPHASGPRRLGCRERAMNHSVREAGVPLGLPAPGARRRGSGSASRS 253

Db 68 QP 69
QY 254 LP 255

RESULT 4
ID US-08-278-729A-23 STANDARD; PRT; 399 AA.
XX xxxxxx
DT
XX
DE Sequence 23, Application US/08278729A

CC Sequence 23, Application US/08278729A
CC Patent No. 5650276
CC GENERAL INFORMATION:
CC APPLICANT: SMART, JOHN
CC APPLICANT: OPPERMANN, HERMAN
CC APPLICANT: OKAYNAK, ENGIN
CC APPLICANT: KUBERASAMPATH, THANGAVEL
CC APPLICANT: RUEGER, DAVID C
CC APPLICANT: PANG, ROY H L
CC APPLICANT: COHEN, CHARLES M
CC TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/278,729A
CC FILING DATE: 20-JUL-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER ESQ, EDMUND R
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-058CPFW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508) 435-9001
CC TELEFAX: (508) 435-6951
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA: 44764 MW: 790568 CN;

Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

QY	254 LP	255
----	--------	-----

CC sequence 23, Application US/08643563A
CC Patent No. 5707810


```
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/155,343A
CC      FILING DATE: 15-NOV-1993
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: FENTON Esq., GILLIAN M.
CC      REGISTRATION NUMBER: 36,508
CC      REFERENCE/DOCKET NUMBER: CRP-067FW
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (617) 248-7560
CC      TELEFAX: (617) 248-7100
CC      INFORMATION FOR SEQ ID NO: 23:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 399 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 399 AA; 44764 MW; 790568 CN;
CC
CC      QUERY MATCH          1.4%; Score 118; DB 1; Length 399;
CC      Best Local Similarity 38.7%; Pred. No. 2,06e+00;
CC      Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;
CC
Dd    12 GLATCATGGGHHPRPP-HTC-PORRLG-ARERRDMORELIATVLGR-PPRAQPAAR 67
Oy    194 GPLPLTQLCAATGARPPHASPSPRRRLGCERAMNNSYREAGVPLGLPALPGARRRGSSRS 253
Db    68 QP 69
Qy    254 LP 255

```

```
DE      RESULT 10
ID     PCT-US93-07231-23 STANDARD; PRI: 399 AA.
XX     xxxxxx
DX
DT
DY
XX     Sequence 23, Application PC/TUS9307231
XX
CC     Sequence 23, Application PC/TUS9307231
CC     GENERAL INFORMATION:
CC     APPLICANT:
CC     TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND
CC     TITLE OF INVENTION: REPAIR
CC     NUMBER OF SEQUENCES: 33
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: CREATIVE BIOMOLECULES, INC.
CC     STREET: 35 SOUTH STREET
CC     CITY: HOPKINTON
CC     STATE: MASSACHUSETTS
CC     COUNTRY: USA
CC     ZIP: 01748
CC     COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: Patent Release #1.0, Version #1.25
CC     CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: PCT/US93/07231
CC     FILING DATE: 19930729
CC     CLASSIFICATION:
CC     ATTORNEY/AGENT INFORMATION:
CC     NAME: KELLEY, ROBIN D.
CC     REGISTRATION NUMBER: 34,637
CC     REFERENCE/DOCKET NUMBER: CRP-070
CC     TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: 617/248-7000
CC
```

CC	TELEFAX: 617/248-7100	DB	Score 118; DB 2; Length 399;
CC	INFORMATION FOR SEQ. ID NO: 23;		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 399 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 399 AA; 44764 MW; 790568 CN;		
CC			
CC	Query Match		
CC	Best Local Similarity 38.7%; Pred. No. 2.06e+00;		
CC	Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;		
CC			
CC	Db		
CC	68 QP 69		
CC	QY 254 LP 255		
CC			
CC	RESULT 11		
CC	PCF-US93-05446-13 STANDARD; PRT; 399 AA.		
CC	XX		
CC	XXXXXX		
CC			
CC	Sequence 13, Application PC/TUS9305446		
CC	DE		
CC	Sequence 13, Application PC/TUS9305446		
CC	GENERAL INFORMATION:		
CC	APPLICANT:		
CC	TITLE OF INVENTION: PROSTHETIC DEVICES HAVING ENHANCED		
CC	TITLE OF INVENTION: OSTEOGENIC PROPERTIES		
CC	NUMBER OF SEQUENCES: 22		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Creative Biomolecules, Inc.		
CC	STREET: 35 South Street		
CC	CITY: Hopkinton		
CC	STATE: MA		
CC	COUNTRY: USA		
CC	ZIP: 01748		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: Patentin Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: PCT/US93/05446		
CC	FILING DATE: 19930608		
CC	CLASSIFICATION:		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: PITCHER ESQ, EDMUND R		
CC	REGISTRATION NUMBER: 27,829		
CC	REFERENCE/DOCKET NUMBER: STK-057		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 617/248-7000		
CC	INFORMATION FOR SEQ. ID NO: 13:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 399 amino acids		
CC	TYPE: AMINO ACID		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
CC	SEQUENCE 399 AA; 44764 MW; 790568 CN;		
CC			
CC	Query Match		
CC	Best Local Similarity 38.7%; Pred. No. 2.06e+00;		
CC	Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;		
CC			
CC	Db		
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQRELLAVGLPGR-PRPRAQPPAAR 67		
CC	QY 194 GPPLYQLGAAQAQAPPHPASGPRRRRLCCERAMHNSVREAGVPLGLPAPGARRRGGSASRS 253		
CC			
CC	12 GLALCALGGGGRPP-HTC-PQRLG-ANRRMDQ		

[illegible]

XX Sequence 23, Application US/08462623
CC Patent No. 5739107
CC GENERAL INFORMATION:
CC APPLICANT: COHEN, CHARLES M.
CC APPLICANT: CHARETTE, MARC F.
CC APPLICANT: KUBERASAMPATH, THANGAVEL
CC APPLICANT: RUEGER, DAVID C.
CC APPLICANT: OPPERMAN, HERMANN
CC APPLICANT: PANG, ROY H.L.
CC APPLICANT: OKRAYNAK, ENGIN
CC APPLICANT: SMART, JOHN E.
CC TITLE OF INVENTION: MORPHOGEN TREATMENT OF GASTROINTESTINAL
CC TITLE OF INVENTION: ULCERS.
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/462,623
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/445,862
CC FILING DATE: 22-MAY-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FENTON Esq., GILLIAN M.
CC REGISTRATION NUMBER: 36,508
CC REFERENCE/DOCKET NUMBER: CRP-074CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508) 435-9901
CC TELEFAX: (508) 435-6951
CC INFORMATION FOR SEQ ID NO: 23:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;
SQ
Query Match 1.4%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;
DB 12 GLAICALGGHGPRPP-HTC-PORRLG-ARERRDMOREILAVIGLGR-PRPRAQPAAR 67
QY 194 GPPLYQLGAATQARPPPHASGPRRLGCEERAMNHSYREAGVPLGLPAPGARRRGSASRS 253
DB 68 QP 69
QY 254 LP 255
RESULT 15
ID PCT-0593-07189-8 STANDARD; PRT; 399 AA.
XX
XX xxxxxx
XX
XX
XX Sequence 8, Application PC/TUS9307189
XX Sequence 8, Application PC/TUS9307189
CC GENERAL INFORMATION:

CC APPLICANT:
CC TITLE OF INVENTION: NOVEL MORPHOGENIC PROTEIN COMPOSITIONS
CC TITLE OF INVENTION: OF MATTER
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
CC STREET: 35 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/07189
CC FILING DATE: 19930729
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: KELLEY, ROBIN D.
CC REGISTRATION NUMBER: 34,637
CC REFERENCE/DOCKET NUMBER: CRP-081CP
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 399 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 399 AA; 44764 MW; 790568 CN;
SQ
Query Match 1.4%; Score 118; DB 2; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.06e+00;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;
DB 12 GLAICALGGHGPRPP-HTC-PORRLG-ARERRDMOREILAVIGLGR-PRPRAQPAAR 67
QY 194 GPPLYQLGAATQARPPPHASGPRRLGCEERAMNHSYREAGVPLGLPAPGARRRGSASRS 253
DB 68 QP 69
QY 254 LP 255
Search completed: Fri Dec 18 18:31:03 1998
Job time : 32 secs.

WIRE
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSPrch_PP protein - protein database search, using Smith-Waterman algorithm

Run on: Fri Dec 18 18:31:23 1998; MasPar time 55.03 Seconds

Tabular output not generated. 793.597 Million cell updates/sec

Title: >US-08-951-733-20

Description: (1-1154) from US08951733.pep

Perfect Score: 8624

Sequence: 1 HASGQRCVLRTWEALAPAT.....TALEAANPALPSDFITLD 1154

Scoring table: PAM 150

Gap 11

Searched: 321025 segs, 3784181 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

a-pending
1:P9 2:U60 3:U7 4:U80 5:U81 6:U82 7:U83 8:U84 9:U85
10:U86 11:U87 12:U88 13:U89 14:U90 15:U91 16:NEWP
17:NEWU6 18:NEWU8 19:NEWU9

Statistics: Mean 40.459; Variance 196.267; scale 0.206

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	8624	100.0	1154	13	US-08-951-Sequence 20, Applicat	0.00e+00
2	8607	99.8	1189	13	US-08-974-Sequence 613, Applicat	0.00e+00
3	8607	99.8	1189	13	US-08-911-Sequence 34, Applicat	0.00e+00
4	8607	99.8	1189	13	US-08-912-Sequence 325, Applicat	0.00e+00
5	8607	99.8	1200	18	US-08-974-Sequence 612, Applicat	0.00e+00
6	8607	99.8	1200	13	US-08-912-Sequence 324, Applicat	0.00e+00
7	8607	99.8	1200	13	US-08-911-Sequence 33, Applicat	0.00e+00
8	8607	99.8	1285	18	US-08-874-Sequence 600, Applicat	0.00e+00
9	8607	99.8	1285	13	US-08-911-Sequence 32, Applicat	0.00e+00
10	8607	99.8	1285	13	US-08-912-Sequence 314, Applicat	0.00e+00
11	8590	99.6	1407	13	US-08-911-Sequence 55, Applicat	0.00e+00
12	8590	99.6	1407	18	US-08-974-Sequence 628, Applicat	0.00e+00
13	8580	99.6	1407	13	US-08-912-Sequence 334, Applicat	0.00e+00
14	8455	98.2	1132	18	US-08-874-Sequence 2, Applicatio	0.00e+00
15	8455	98.2	1132	13	US-08-912-Sequence 2, Applicatio	0.00e+00
16	8455	98.2	1132	13	US-08-911-Sequence 2, Applicatio	0.00e+00
17	8455	98.2	1132	14	US-09-042-Sequence 3, Applicatio	0.00e+00
18	8455	98.2	1132	12	US-09-052-Sequence 2, Applicatio	0.00e+00
19	8455	98.2	1132	14	US-08-854-Sequence 225, Applicat	0.00e+00
20	8455	98.2	1154	13	US-08-912-Sequence 323, Applicat	0.00e+00
21	8465	98.2	1154	18	US-08-974-Sequence 611, Applicat	0.00e+00

RESULT ID	1	US-08-951-733-20	STANDARD:	PRT:	1154 AA.
XX	US-08-951-733-20	xxxxxx			
XX	Sequence 20, Application US/08951733				
DE	Sequence 20, Application US/08951733				
CC	Sequence 20, Application US/08951733				
CC	GENERAL INFORMATION:				
CC	APPLICANT: Harrington, Lea A.				
CC	APPLICANT: Robinson, Murray O.				
CC	TITLE OF INVENTION: NOVEL GENES ENCODING TELOMERASE PROTEINS				
CC	NUMBER OF SEQUENCES: 44				
CC	CORRESPONDENCE ADDRESS:				
CC	ADDRESSEE: Amgen Inc.				
CC	STREET: One Amgen Center Drive				
CC	CITY: Thousand Oaks				
CC	STATE: CA				
CC	COUNTRY: USA				
CC	ZIP: 91320-1789				
CC	COMPUTER READABLE FORM:				
CC	MEDIUM TYPE: Floppy disk				
CC	COMPUTER: IBM PC compatible				
CC	OPERATING SYSTEM: PC-DOS/MS-DOS				
CC	SOFTWARE: Patent Release #1.0, Version #1.30				
CC	CURRENT APPLICATION DATA:				
CC	APPLICATION NUMBER: US/08/951,733				
CC	FILING DATE: 16-OCT-1997				
CC	CLASSIFICATION: 435				
CC	PRIOR APPLICATION DATA:				
CC	APPLICATION NUMBER: US 08/873,039				
CC	FILING DATE: 11-JUN-1997				
CC	PRIOR APPLICATION DATA:				
CC	APPLICATION NUMBER: US 08/751,189				
CC	FILING DATE: 15-NOV-1996				
CC	ATTORNEY/AGENT INFORMATION:				
CC	NAME: Oleski, Nancy A.				
CC	REGISTRATION NUMBER: 34,688				
CC	REFERENCE/DOCKET NUMBER: A-433B				
CC	TELECOMMUNICATION INFORMATION:				
CC	TELEPHONE: (805) 447-6504				
CC	TELEFAX: (805) 499-8011				
CC	INFORMATION FOR SEQ ID NO: 20:				

SEQUENCE CHARACTERISTICS:
LENGTH: 1154 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
SEQUENCE 1154 AA; 129326 MW; 6842246 CN;

Query Match 100.0%; Score 8624; DB 13; Length 1154;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1154; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 HASGORCVLTWEALAPATPAMPAPRCRAVRSLLRSHREYVPLATFVRRLGPOGRL 60
QY 1 HASGORCVLTWEALAPATPAMPAPRCRAVRSLLRSHREYVPLATFVRRLGPOGRL 60
Db 61 VORGDPAFAFALVAQCIVCPMDARPPAPASFRVOSCLKEVLAVRLORLCEGAKNYLA 120
QY 61 VORGDPAFAFALVAQCIVCPMDARPPAPASFRVOSCLKEVLAVRLORLCEGAKNYLA 120
Db 121 FGFALLDAGRGPPPEAFSTVSRYLPTNTVDALRSGAMGLLRVGDVTVHLIARCAL 180
QY 121 FGFALLDAGRGPPPEAFSTVSRYLPTNTVDALRSGAMGLLRVGDVTVHLIARCAL 180
Db 181 FVLVAAPSCAYVCGPPLVQLGAATQARPPHASGPRRLGGERAMNSVREAGVPLGLPA 240
QY 181 FVLVAAPSCAYVCGPPLVQLGAATQARPPHASGPRRLGGERAMNSVREAGVPLGLPA 240
Db 241 PGARRRGSSASRSLPLPRPRRGAPPEPERTVVGSGSWAHGRTGSPDSRGCVVSPAR 300
QY 241 PGARRRGSSASRSLPLPRPRRGAPPEPERTVVGSGSWAHGRTGSPDSRGCVVSPAR 300
Db 301 ABEATSLGALSGTRHSHPSVSGROHHAAGPSTSPRPWDTPCPPVYAEYTHFLYSDDK 360
QY 301 ABEATSLGALSGTRHSHPSVSGROHHAAGPSTSPRPWDTPCPPVYAEYTHFLYSDDK 360
Db 361 EOLRPSFLLSRLPSLGTGARLVEITFI GSRPMWGTGTRRLPLRQRMORPLFELTIG 420
QY 361 EOLRPSFLLSRLPSLGTGARLVEITFI GSRPMWGTGTRRLPLRQRMORPLFELTIG 420
Db 421 NHAOCPYGLVLTCHPLRAAVTPAAGVCAREKPOGSAVAPEEEDTPRRLVQLRQHSPP 480
QY 421 NHAOCPYGLVLTCHPLRAAVTPAAGVCAREKPOGSAVAPEEEDTPRRLVQLRQHSPP 480
Db 481 WQYGFVACLRRLVPCGLMGRHNERPLNNTKFFISLGHAKLSLOELTWKMSVRCOA 540
QY 481 WQYGFVACLRRLVPCGLMGRHNERPLNNTKFFISLGHAKLSLOELTWKMSVRCOA 540
Db 541 WLRRSPGVGCPAAEHRLREELIAKFLHMLMSVYVELLSRFEYVETTFQKNRLEFYRK 600
QY 541 WLRRSPGVGCPAAEHRLREELIAKFLHMLMSVYVELLSRFEYVETTFQKNRLEFYRK 600
Db 601 SVWSKLSIGIRHOLKRVOLRELSAEVYRQHEARAPALLSTRLEPIPRPDGLRIVNNDY 660
QY 601 SVWSKLSIGIRHOLKRVOLRELSAEVYRQHEARAPALLSTRLEPIPRPDGLRIVNNDY 660
Db 661 VVGARTFERREKRAERLTSRVKALFSLVNERARRPGLLGASVLDGDDIHRAMRFEVLVR 720
QY 661 VVGARTFERREKRAERLTSRVKALFSLVNERARRPGLLGASVLDGDDIHRAMRFEVLVR 720
Db 721 AODPPPELYFVKVDVTGAVDTIPORLCEVLAIIKPNQTCVRRYAVVQAAHGHVKA 780
QY 721 AODPPPELYFVKVDVTGAVDTIPORLCEVLAIIKPNQTCVRRYAVVQAAHGHVKA 780
Db 781 FKSHTSLTDLOPYRKQVVAHLQETSPRLDAVYIEOSSLSNEASSGLDVEFLRECHAV 840
QY 781 FKSHTSLTDLOPYRKQVVAHLQETSPRLDAVYIEOSSLSNEASSGLDVEFLRECHAV 840
Db 841 RIRKSYVQCGIPIOGSILSTLCSLCYGDMEKFLAGIRDGILLRLVDDFLVTLPHLT 900
QY 841 RIRKSYVQCGIPIOGSILSTLCSLCYGDMEKFLAGIRDGILLRLVDDFLVTLPHLT 900
Db 901 HAKTFLRLTVRGVPEYGCVMNLKTVVNFVYDEALGTAIVQMPAHGLFPWCGILLDTR 960

|||||
QY 901 HAKTFLRLTVRGVPEYGCVMNLKTVVNFVYDEALGTAIVQMPAHGLFPWCGILLDTR 960
Db 961 TLEVQSDYSSTARTSTRASTENRGRAGNMRKLPGLVRLKCHSLFDLOVNSIQIYC 1020
QY 961 TLEVQSDYSSTARTSTRASTENRGRAGNMRKLPGLVRLKCHSLFDLOVNSIQIYC 1020
Db 1021 TNYIKILLQAYRPHACVLOLPEHQOVMKMPTEFLVVISPTASTLCYSILKAKNAGSIGA 1080
QY 1021 TNYIKILLQAYRPHACVLOLPEHQOVMKMPTEFLVVISPTASTLCYSILKAKNAGSIGA 1080
Db 1081 KGAAGPLPSEAVOMLCHOAFLKLTTRHYVYVPLDGLSLRTAOTOLSRLKPLGTTLTALAEA 1140
QY 1081 KGAAGPLPSEAVOMLCHOAFLKLTTRHYVYVPLDGLSLRTAOTOLSRLKPLGTTLTALAEA 1140
Db 1141 ANPALPSEDFKTIID 1154
QY 1141 ANPALPSEDFKTIID 1154

RESULT 2
ID US-08-974-549-613 STANDARD: PRT: 1189 AA.
AC xxxxxx
DT
XX
XX
DE Sequence 613, Application US/08974549
CC Sequence 613, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cecch, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morlin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 05-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312

CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Red
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 613:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1189 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1189
CC OTHER INFORMATION: /note= "fusion protein composed of
CC OTHER INFORMATION: melittin signal sequence and full length
CC OTHER INFORMATION: hTERT protein"
CC SEQUENCE 1189 AA: 133179 MW: 7256545 CN:

Query Match 99.8%; Score 8607; DB 18; Length 1189;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db 37 ASTORCVLLRTWELALPATPAMPBRADRCRAVRSLSRSHYREVLPPLATFYRRLLGPGQWRVLY 96
Qy 2 ASGQRCLVLLRTWELALPATPAMPBRADRCRAVRSLSRSHYREVLPPLATFYRRLLGPGQWRVLY 61
Db 97 QGSDPAFAFALVAQCVCVPMQARPPAPPSFQVQSCLELVARVYQRLCERAKNVLA 156
Qy 62 QRGDPAFAFALVAQCVCVPMQARPPAPPSFQVQSCLELVARVYQRLCERAKNVLA 121
Db 157 GFALLDGGAGGPEAFTTSVRSYLPMTVTDALRGSGAMGLLRVGGDDVLYHLLARCALE 216
Qy 122 GFALLDGGAGGPEAFTTSVRSYLPMTVTDALRGSGAMGLLRVGGDDVLYHLLARCALE 181
Db 217 VLVAPSCAYOVCGPPLYOLGATQARPPPHAGSPRRRLGCERAMNHSVREAGVPLGLPAP 276
Qy 182 VLVAPSCAYOVCGPPLYOLGATQARPPPHAGSPRRRLGCERAMNHSVREAGVPLGLPAP 241
Db 277 GARRRGGSASRSPLPKRRRGAAPPERTPVQSGSNABRGTRGSDGFCVVSAPARA 336
Qy 242 GARRRGGSASRSPLPKRRRGAAPPERTPVQSGSNABRGTRGSDGFCVVSAPARA 301
Db 337 BEATSLGALSGTRHSHPSGROHAGPPSTSRPPMDPCPPVYAETKHFLLSYSGDKE 396
Qy 302 BEATSLGALSGTRHSHPSGROHAGPPSTSRPPMDPCPPVYAETKHFLLSYSGDKE 361
Db 397 QLRPSFLLSLRPSLTGARLVETIFLGSRPWMPGTPRLLPRLPQRYWQMRPLLELLGN 456
Qy 362 QLRPSFLLSLRPSLTGARLVETIFLGSRPWMPGTPRLLPRLPQRYWQMRPLLELLGN 421
Db 457 HAOCPPVLLKTCPLRAAVTPAAGVCARPKPGSVAAPEEEDTDPRRLVQLLRHSSPW 516
Qy 422 HAOCPPVLLKTCPLRAAVTPAAGVCARPKPGSVAAPEEEDTDPRRLVQLLRHSSPW 481
Db 517 QVAGFVRACLRRLVPPGLMGSRHNERFLNKKFTSLGHAFTLSLOELTWKMSVDCAM 576
Qy 517 QVAGFVRACLRRLVPPGLMGSRHNERFLNKKFTSLGHAFTLSLOELTWKMSVDCAM 576

Qy 482 QVAGFVRACLRRLVPPGLMGSRHNERFLNKKFTSLGHAFTLSLOELTWKMSVDCAM 541
Db 577 LRRSPGVCVPAEHRLEBEILAKFLHMLMSYVVELLRSPFYVETEFQKRLPEYRS 636
Qy 542 LRRSPGVCVPAEHRLEBEILAKFLHMLMSYVVELLRSPFYVETEFQKRLPEYRS 601
Db 637 VMSKQSIGIRQHLKRVOLRELSAEVROHREARPALTSRLRFIPKPDGLRPYVMDYV 696
Qy 602 VMSKQSIGIRQHLKRVOLRELSAEVROHREARPALTSRLRFIPKPDGLRPYVMDYV 661
Db 697 VGARTFRREKRAERLTSRYKALFSVLTNERARBPGLGASVGLDDIHRAMTFVLYVRA 756
Qy 662 VGARTFRREKRAERLTSRYKALFSVLTNERARBPGLGASVGLDDIHRAMTFVLYVRA 721
Db 757 QPPPELVYKVDYMGAVDTIPQDRLTEVIAIIPQNTYCVRRAYVOKAHGVRAE 816
Qy 722 QPPPELVYKVDYMGAVDTIPQDRLTEVIAIIPQNTYCVRRAYVOKAHGVRAE 781
Db 817 KSHVSTLFDLPYMRQFVAHLOETSPLRDAVYIEOSSSLNEASSGLFDVFLRFMCHAVR 876
Qy 782 KSHVSTLFDLPYMRQFVAHLOETSPLRDAVYIEOSSSLNEASSGLFDVFLRFMCHAVR 841
Db 877 IRGKSYVQCQIGIPQGSILSTLLCSLCYGDMEKRLFAGIRRDGLLRVDDFLVTPHLLTH 936
Qy 842 IRGKSYVQCQIGIPQGSILSTLLCSLCYGDMEKRLFAGIRRDGLLRVDDFLVTPHLLTH 901
Db 937 AKTFELTLVRYGPEYGCYVNLAKTYVNFVEDEALGTAFOVMPAHGLEPWCGLLDRT 996
Qy 902 AKTFELTLVRYGPEYGCYVNLAKTYVNFVEDEALGTAFOVMPAHGLEPWCGLLDRT 961
Db 997 LEVQSDYSYATSIASLTFNRFGRAGNMNRKLFGLVRLKCHSLFDLDQVNSLOTYCT 1056
Qy 962 LEVQSDYSYATSIASLTFNRFGRAGNMNRKLFGLVRLKCHSLFDLDQVNSLOTYCT 1021
Db 1057 NIYKILLQAYRFAVCYQLPPEHQVWKNPTEFLRYISDPAISLCAKSNAGSLGAK 1116
Qy 1022 NIYKILLQAYRFAVCYQLPPEHQVWKNPTEFLRYISDPAISLCAKSNAGSLGAK 1081
Db 1117 GAAGPLPSEAVOMLCHOAFLKLTIRHRTYVPLLSLRPAQVQLSRKLPGLTLLALEAA 1176
Qy 1082 GAAGPLPSEAVOMLCHOAFLKLTIRHRTYVPLLSLRPAQVQLSRKLPGLTLLALEAA 1141
Db 1177 NPALPSDFKTIID 1189
Qy 1142 NPALPSDFKTIID 1154

RESULT 3
ID US-08-911-312-34 STANDARD; PRT: 1189 AA.
AC xxxxxx
DT
XX
XX
DE Sequence 34, Application US/08911312
CC Sequence 34, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA

```

CC      ZIP: 94111-3834
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patentin Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/911,312
CC      FILING DATE: 14-AUG-1997
CC      CLASSIFICATION: 536
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/724,643
CC      FILING DATE: 01-OCT-1996
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/844,419
CC      FILING DATE: 18-APR-1997
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/846,017
CC      FILING DATE: 25-APR-1997
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/851,843
CC      FILING DATE: 06-MAY-1997
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/854,050
CC      FILING DATE: 09-MAY-1997
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/912,951
CC      FILING DATE: 14-AUG-1997
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/915,503
CC      FILING DATE: 14-AUG-1997
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Elmhorn, Gregory P.
CC      REGISTRATION NUMBER: 36,440
CC      REFERENCE/DOCKET NUMBER: 015389-002500US
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (415) 576-0200
CC      TELEFAX: (415) 576-0300
CC      INFORMATION FOR SEQ ID NO: 34:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 1189 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS:
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 1189 AA; 133179 MW; 7256545 CN;
SQ
Query Match      99.8%; Score 8607; DB 13; Length 1189;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      302  EATSLGALSGRHSHSPVGRQHAGPPSTSRPPRMDTPCPVYAETKHFLYSSGDKE 361
DB      397  QLRSPFLSSLRPSLTGARRLVETIFLGSRRPMGTRRLPLRQRTWQNRPLLELLGN 456
QY      362  QLRSPFLSSLRPSLTGARRLVETIFLGSRRPMGTRRLPLRQRTWQNRPLLELLGN 421
DB      457  HAQCPYGLTKHCPPLRAAVTPAAGVCAREKPOGGSVAPEEEDDPRLVQLLRQHSPP 516
QY      422  HAQCPYGLTKHCPPLRAAVTPAAGVCAREKPOGGSVAPEEEDDPRLVQLLRQHSPP 481
DB      517  QVYGFVRACLRRLVPPGLWSRHNERFLRNTKFFISLGHAKLSLOELTKMSVRDCAW 576
QY      482  QVYGFVRACLRRLVPPGLWSRHNERFLRNTKFFISLGHAKLSLOELTKMSVRDCAW 541
DB      577  LRRSPGVGCVAHAHRLEELTFLWMSVYVELLSFFYTETTPQKNRFFYRKS 636
QY      542  LRRSPGVGCVAHAHRLEELTFLWMSVYVELLSFFYTETTPQKNRFFYRKS 601
DB      637  VMSKLQSIGIRQHLKRVQLRELSEAEYRQREARPALITSRLRFPDGLRPIVNDYV 696
QY      602  VMSKLQSIGIRQHLKRVQLRELSEAEYRQREARPALITSRLRFPDGLRPIVNDYV 661
DB      697  VGARTFRERKRAERLTSRVKALFSVLNTERARBPGLGASVGLDDIHRAMRFEVLRYA 756
QY      662  VGARTFRERKRAERLTSRVKALFSVLNTERARBPGLGASVGLDDIHRAMRFEVLRYA 721
DB      757  QDPPELTFYKAVDYTGAYDITPDRLTEVASTIKPONTICVRRYAVYQRAAHGVKAF 816
QY      722  QDPPELTFYKAVDYTGAYDITPDRLTEVASTIKPONTICVRRYAVYQRAAHGVKAF 781
DB      817  KSHVSTLTDLPYKQFVAHLQETSPLRDVAIVIOSSSLNEASGLPDVFLREWCNHAVR 876
QY      782  KSHVSTLTDLPYKQFVAHLQETSPLRDVAIVIOSSSLNEASGLPDVFLREWCNHAVR 841
DB      877  IRGKSYVOCQGIPOGSLSTLSCYGDMEENKLFAGIRBDGLLRVLDVFLVTPHLTH 936
QY      842  IRGKSYVOCQGIPOGSLSTLSCYGDMEENKLFAGIRBDGLLRVLDVFLVTPHLTH 901
DB      937  AKTLRLTLVRGVPEYGCVMNLRKTVNFPVEDEALGSTAFVQMAHGLFPGCGILLDTRT 996
QY      902  AKTLRLTLVRGVPEYGCVMNLRKTVNFPVEDEALGSTAFVQMAHGLFPGCGILLDTRT 961
DB      997  LEVOSDYSYARTSIRASLTFNRFKAGRNMRKLFGLRLKCHSLDLDOVNSLQTVCT 1056
QY      962  LEVOSDYSYARTSIRASLTFNRFKAGRNMRKLFGLRLKCHSLDLDOVNSLQTVCT 1021
DB      1057  NIKYILLQAYRFHACVQLQPFHQOVKNPTFFLRVVISDTASLICYSLKAKNAGMSLGAK 1116
QY      1022  NIKYILLQAYRFHACVQLQPFHQOVKNPTFFLRVVISDTASLICYSLKAKNAGMSLGAK 1081
DB      1117  GAAGPLPSEANQVLTCHQAFLLKLRHRYTVYPLLSGSLRTAQTQSLRKLPGTTTLALPAA 1176
QY      1082  GAAGPLPSEANQVLTCHQAFLLKLRHRYTVYPLLSGSLRTAQTQSLRKLPGTTTLALPAA 1141
DB      1177  NPALPSPDEKTIID 1189
QY      1142  NPALPSPDEKTIID 1154

```

```

RESULT      4
ID          US-08-912-951-325      STANDARD;      PRT;      1189 AA.
AC          xxxxxx
XX          xxxxxx
DT          xxxxxx
XX          xxxxxx
DE          Sequence 325, Application US/08912951
CC          GENERAL INFORMATION:
CC          APPLICANT: Cech, Thomas R.
CC          APPLICANT: Lingner, Joachim
CC          APPLICANT: Nakamura, Toru

```

CC APPLICANT: Chapman, Karen B.
 CC APPLICANT: Morin, Gregg B.
 CC APPLICANT: Harley, Calvin
 CC APPLICANT: Andrews, William H.
 CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
 CC NUMBER OF SEQUENCES: 335
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Townsend and Townsend and Crew LLP
 CC STREET: Two Embarcadero Center, 8th Floor
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: United States of America
 CC ZIP: 94111
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/912,951
 CC FILING DATE: 14-AUG-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/854,050
 CC FILING DATE: 09-MAY-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/851,843
 CC FILING DATE: 06-MAY-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/846,017
 CC FILING DATE: 25-APR-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/844,419
 CC FILING DATE: 18-APR-1997
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/724,643
 CC FILING DATE: 01-OCT-1996
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Apple, Randolph T.
 CC REGISTRATION NUMBER: 36,429
 CC REFERENCE/DOCKET NUMBER: 015389-002600US
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (415) 576-0300
 CC TELEFAX: (415) 576-0300
 CC INFORMATION FOR SEQ ID NO: 325:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 1189 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS:
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 1189 AA; 133179 MW; 7256545 CN;
 SQ
 Query Match 99.8%; Score 8607; DB 13; Length 1189;
 Best Local Similarity 99.8%; Pred. No. 0.00e+00;
 Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 122 GFALLDGAARGGPEAFETTSVRSYLPNTVTDALRGSGAMGLLRVGGDVLVHLRCALE 181
 DB VLVAPSCAYVCYCPPIVOLAATQARPPHSGPRRLGCEBANNHVSREAGVPLGAP 276
 QY 182 VLVAPSCAYVCYCPPIVOLAATQARPPHSGPRRLGCEBANNHVSREAGVPLGAP 241
 DB 277 GARRRGGSASRSLPLPKPRRGAAPPEPERTVGGOSMAHPERTGSPDRGCVVSPAPA 336
 QY 242 GARRRGGSASRSLPLPKPRRGAAPPEPERTVGGOSMAHPERTGSPDRGCVVSPAPA 301
 DB 337 EEAATSLGALSSTNRSHSBSVGRQHAGBPSTSRPPMPDTPCPVYATKHFILYSSGDK 396
 QY 302 EEAATSLGALSSTNRSHSBSVGRQHAGBPSTSRPPMPDTPCPVYATKHFILYSSGDK 361
 DB 397 QLRPSFLLSSLRPSLTGARRLVETIFLGRPMWMTGTPRRLRLQRYWQMPLLELLGN 456
 QY 362 QLRPSFLLSSLRPSLTGARRLVETIFLGRPMWMTGTPRRLRLQRYWQMPLLELLGN 421
 DB 457 HAOCYGVLTGHPPLRAAVTPAAGVCAREKPOGSVAAPEDDTPRRLVOLLRQHSPPM 516
 QY 422 HAOCYGVLTGHPPLRAAVTPAAGVCAREKPOGSVAAPEDDTPRRLVOLLRQHSPPM 481
 DB 517 QYGFVRACLRLVPPGGLMGSRHNERPLRNTKFFISLGHAKUSLOELTWKMSVRCDAW 576
 QY 482 QYGFVRACLRLVPPGGLMGSRHNERPLRNTKFFISLGHAKUSLOELTWKMSVRCDAW 541
 DB 577 LRRSPGVCVPAAEHRLAEELIAFLHMLSVYVELLSRFYETTFQKNRLEFFYKKS 636
 QY 542 LRRSPGVCVPAAEHRLAEELIAFLHMLSVYVELLSRFYETTFQKNRLEFFYKKS 601
 DB 637 VMSKLOSIGICHOHLKRVOLRELSAEVQHQREARPAALTSRLFPKPDGRPIYNDYV 696
 QY 602 VMSKLOSIGICHOHLKRVOLRELSAEVQHQREARPAALTSRLFPKPDGRPIYNDYV 661
 DB 697 VGARTFRERKRAERLTSVKALFVLYNERARBPGLGASVGLDDIHRAWTEFLVRYA 756
 QY 662 VGARTFRERKRAERLTSVKALFVLYNERARBPGLGASVGLDDIHRAWTEFLVRYA 721
 DB 757 QDPPELTFVYVDYTGADTIPODRLTEVIASIIKPOUTCYVRKRAYVQKAHGHVRAAF 816
 QY 722 QDPPELTFVYVDYTGADTIPODRLTEVIASIIKPOUTCYVRKRAYVQKAHGHVRAAF 781
 DB 817 KSHYSTLTDLQPYMKQFAHLOETSPRLDAVYIEOSSSLNASSGLFDPVFLRFMCHHAVR 876
 QY 782 KSHYSTLTDLQPYMKQFAHLOETSPRLDAVYIEOSSSLNASSGLFDPVFLRFMCHHAVR 841
 DB 877 IRGKSYVOCGIPQGSILSTLCSLCYGDMEKMLFAGIRBDGLLRVLDLVTPLHLTH 936
 QY 842 IRGKSYVOCGIPQGSILSTLCSLCYGDMEKMLFAGIRBDGLLRVLDLVTPLHLTH 901
 DB 937 AKTEFLRLTVRGVPEYGVNLRKTYVNEPVEDEALGTAIFYQMPAHGLFPWCGLLDTRT 996
 QY 902 AKTEFLRLTVRGVPEYGVNLRKTYVNEPVEDEALGTAIFYQMPAHGLFPWCGLLDTRT 961
 DB 997 LEVQSDYSVAARTSTRASTENRFGKAGRNMRKTLFGVLRKCHSLFEDLVNLSLOTCT 1056
 QY 962 LEVQSDYSVAARTSTRASTENRFGKAGRNMRKTLFGVLRKCHSLFEDLVNLSLOTCT 1021
 DB 1057 NIKYKILLQAFRFAVCYQLPFOQVWKNPTEFLRVISDTASLCYSILKANAGMSLGAK 1116
 QY 1022 NIKYKILLQAFRFAVCYQLPFOQVWKNPTEFLRVISDTASLCYSILKANAGMSLGAK 1081
 DB 1117 GAAGPLPSEAVQWOLCHQAFLLKTRHRVTYVPLUGSLRTAOTOLSRLKIPGTTTLALEAAA 1176
 QY 1082 GAAGPLPSEAVQWOLCHQAFLLKTRHRVTYVPLUGSLRTAOTOLSRLKIPGTTTLALEAAA 1141
 DB 1177 NPALPSDEKTLID 1189
 QY 1142 NPALPSDEKTLID 1154
 RESULT 5
 ID US-08-974-549-612 STANDARD; PRT; 1200 AA.

XX xxxxxx
AC
XX
DT
XX
Sequence 612, Application US/08974549
XX
Sequence 612, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morlin, Gregg B.
CC APPLICANT: Hartley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17865
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 612:
CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 1200 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1200
CC OTHER INFORMATION: /note="fusion protein composed of His6
CC OTHER INFORMATION: and Anti-Xpress tags, enterokinase
CC OTHER INFORMATION: cleavage site and full length hprt
CC OTHER INFORMATION: protein"
CC SEQUENCE 1200 AA: 134322 MW: 7387257 CN:
SQ
Query Match 99.88; Score 8607; DB 18; Length 1200;
Best Local Similarity 99.98; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 48 ASTGRCVLTWELAPATPAMPAPRCRAVRSILRSYREVLPATFVRLAPGQWRV 107
|||
Qy 2 ASGRCVLTWELAPATPAMPAPRCRAVRSILRSYREVLPATFVRLAPGQWRV 61
Db 108 QRGDPAFRLVAQCLVCPWMDAPPAPSPFROVSCLELVARVYLQRLCEGAKNYLAF 167
|||
Qy 62 QRGDPAFRLVAQCLVCPWMDAPPAPSPFROVSCLELVARVYLQRLCEGAKNYLAF 121
Db 168 GFALIDGARGPPPAFTTSTYSLPNTVTDALNGSGAMGILLRVGDDVYLHLARCALF 227
|||
Qy 122 GFALIDGARGPPPAFTTSTYSLPNTVTDALNGSGAMGILLRVGDDVYLHLARCALF 181
Db 228 VLVAPSCAYOVCGPPLVQLAATAQARPSPHAGSPRRRLGGERAMNHSVRAGVPLGLPAP 287
|||
Qy 182 VLVAPSCAYOVCGPPLVQLAATAQARPSPHAGSPRRRLGGERAMNHSVRAGVPLGLPAP 241
Db 288 GARRRGASASRLPLPRPRRGAAPPEERTPVQGSVAHPDRTGSDRGFCVSPARPA 347
|||
Qy 242 GARRRGASASRLPLPRPRRGAAPPEERTPVQGSVAHPDRTGSDRGFCVSPARPA 301
Db 348 EATSLGALSGTRHSHSPSVGRHHAGPSTSRPPRWDTPCCPPVYAEYHFLYSSDKE 407
|||
Qy 302 EATSLGALSGTRHSHSPSVGRHHAGPSTSRPPRWDTPCCPPVYAEYHFLYSSDKE 361
Db 408 QLRSPFLSSLRPSLTGARLVETIFLGSRRPMGPRRLPRPQRWQMRPFLELLGN 467
|||
Qy 362 QLRSPFLSSLRPSLTGARLVETIFLGSRRPMGPRRLPRPQRWQMRPFLELLGN 421
Db 468 HAQCPYVLLKTHCPRLRAAVTPAAGVCAREKPOGSVAAPBEDTDPRLVOLLROHSSPW 527
|||
Qy 422 HAQCPYVLLKTHCPRLRAAVTPAAGVCAREKPOGSVAAPBEDTDPRLVOLLROHSSPW 481
Db 528 QYGFVYACLRRLVPPGSLGSRHNERFLNRTKFTISLGHAKLSLOELTWKSYVRCAN 587
|||
Qy 482 QYGFVYACLRRLVPPGSLGSRHNERFLNRTKFTISLGHAKLSLOELTWKSYVRCAN 541
Db 588 LRSPPGVGCVPAEHRRLREELTLAKFLWLSYVVELLSRFYVTEFTFOKNLFFYRKS 647
|||
Qy 542 LRSPPGVGCVPAEHRRLREELTLAKFLWLSYVVELLSRFYVTEFTFOKNLFFYRKS 601
Db 648 VWSKLSIGIRHILKRVQLRELSEAYRQREARPALTSRLFTIPKDLRPIVNNDDV 707
|||
Qy 602 VWSKLSIGIRHILKRVQLRELSEAYRQREARPALTSRLFTIPKDLRPIVNNDDV 661
Db 708 VGATFERREKRAERLTSRVKALFSVLYTEBARRRPGLIGASVGLGDIHRAWRFTVLRVA 767
|||
Qy 662 VGATFERREKRAERLTSRVKALFSVLYTEBARRRPGLIGASVGLGDIHRAWRFTVLRVA 721
Db 768 QDPPELYEFKVDVTGAVDTIPDRLTEVYASIIKPPONTCVRRYAVYQAAAGHYRKAF 827
|||
Qy 722 QDPPELYEFKVDVTGAVDTIPDRLTEVYASIIKPPONTCVRRYAVYQAAAGHYRKAF 781
Db 828 KSHVSTLTDLPYWRQFVAHLOETSPLRDAVVEQSSSLNEASSGLDFVFLRPMCHHAVR 887
|||
Qy 782 KSHVSTLTDLPYWRQFVAHLOETSPLRDAVVEQSSSLNEASSGLDFVFLRPMCHHAVR 841

DB 888 IRGSYVOCGIPGSIITLLCSLCYGDMMENKLEAGIRRDGLLRVDLWPHLTH 947
CC
CC
QY 842 IRGSYVOCGIPGSIITLLCSLCYGDMMENKLEAGIRRDGLLRVDLWPHLTH 901
CC
DB 948 AKTFLRTVRCVPEVGCYNLRKTYVNFVDEALGTAFAVOMPAHGLFPMCGLLDTRT 1007
CC
QY 902 AKTFLRTVRCVPEVGCYNLRKTYVNFVDEALGTAFAVOMPAHGLFPMCGLLDTRT 961
CC
DB 1008 LEVOSDYSSYARTSIRASLFTNRFKAGRNMRKLEGVLRKCHSLFLDLQVNSLQTVCT 1067
CC
QY 962 LEVOSDYSSYARTSIRASLFTNRFKAGRNMRKLEGVLRKCHSLFLDLQVNSLQTVCT 1021
CC
DB 1068 NIKYLLIQAARFACVQQLPFPHOQVKNRPTFLAVISDTASLCYSIIKANAGSLGAK 1127
CC
QY 1022 NIKYLLIQAARFACVQQLPFPHOQVKNRPTFLAVISDTASLCYSIIKANAGSLGAK 1081
CC
DB 1128 GAAGPLPSEAVOMLCHQAFLLKTRHRTYVPLGSLRTAQTOLSRKLPGLTTLALEAA 1187
CC
QY 1082 GAAGPLPSEAVOMLCHQAFLLKTRHRTYVPLGSLRTAQTOLSRKLPGLTTLALEAA 1141
CC
DB 1188 NPALPSEDKTILD 1200
CC
QY 1142 NPALPSEDKTILD 1154
CC
RESULT 6
ID US-08-912-951-324 STANDARD; PRT; 1200 AA.
AC
XX xxxxxx
XX
DE Sequence 324, Application US/08912951
XX
CC GENERAL INFORMATION:
CC
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morlin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017

CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-00260005
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 324:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1200 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1200 AA; 134322 MW; 7387257 CN;
SQ
Query Match 99.8%; Score 8607; DB 13; Length 1200;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 48 ASTORCVLRTWELALPATPMAPRAAPRAVRSILRSYHREVLPLATFVRRLSGPGRRLV 107
CC
QY 2 ASGRCVLTWELALPATPMAPRAAPRAVRSILRSYHREVLPLATFVRRLSGPGRRLV 61
CC
DB 108 QRGDPAFRAVLAOCVCPWDARPPAAPSFRQVSLKELAVAVLQRLCERGAKNVLA 167
CC
QY 62 QRGDPAFRAVLAOCVCPWDARPPAAPSFRQVSLKELAVAVLQRLCERGAKNVLA 121
CC
DB 168 GFALLDARGGPPPAFTTSVRSYLPNTVTDALRGSGAMGLLRVGDVYLHLLARCALF 227
CC
QY 122 GFALLDARGGPPPAFTTSVRSYLPNTVTDALRGSGAMGLLRVGDVYLHLLARCALF 181
CC
DB 228 VLVAAPSCAYQVCGPPPLVQLGAATQARPPPHASGPRRLGCEPRAVNSVREAGVPLGAP 287
CC
QY 182 VLVAAPSCAYQVCGPPPLVQLGAATQARPPPHASGPRRLGCEPRAVNSVREAGVPLGAP 241
CC
DB 288 GARRRGASASRLPLPRKPRGAAPPEPERPVGGGSAHPGRTGSPDRGCVVSPARA 347
CC
QY 242 GARRRGASASRLPLPRKPRGAAPPEPERPVGGGSAHPGRTGSPDRGCVVSPARA 301
CC
DB 348 EEAATSLGALSGRHSHPVSGRHHAGPPSTSRPPRMDTPCPPVYAEKHFLLYSSGDK 407
CC
QY 302 EEAATSLGALSGRHSHPVSGRHHAGPPSTSRPPRMDTPCPPVYAEKHFLLYSSGDK 361
CC
DB 408 QLRPSFLISSLRPSLTGARLVETIFLGSRPWMPGTRRLRLQRTWQMPRLLELLGN 467
CC
QY 362 QLRPSFLISSLRPSLTGARLVETIFLGSRPWMPGTRRLRLQRTWQMPRLLELLGN 421
CC
DB 468 HAQCPYGVLTHTCPRLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLRQHSPPW 527
CC
QY 422 HAQCPYGVLTHTCPRLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLRQHSPPW 481
CC
DB 528 QYGFVRACLRVLVPGIWSRHNERRFLRNTKFFISLGRHAKLSLOELTWKMSVRCDAW 587
CC
QY 482 QYGFVRACLRVLVPGIWSRHNERRFLRNTKFFISLGRHAKLSLOELTWKMSVRCDAW 541
CC
DB 588 LRSPGVGCVAHAHRRLREETLAFELHMLMSVYVELLRSEFYETETFOKNRLFEYRKS 647
CC
QY 542 LRSPGVGCVAHAHRRLREETLAFELHMLMSVYVELLRSEFYETETFOKNRLFEYRKS 601
CC
DB 648 VMSKLSIGIGIQLHKKRVOLRELSEAEYVQRHREARPALITSLRIPRDPGLRPIVNDYV 707
CC
QY 602 VMSKLSIGIGIQLHKKRVOLRELSEAEYVQRHREARPALITSLRIPRDPGLRPIVNDYV 661
CC

DB 528 QYGVFVACLRRLVPGILMGSRHNRRLRNKFTSLGKNAKLSIQELTWMKSYRDCAW 587
OY 482 QYGVFVACLRRLVPGILMGSRHNRRLRNKFTSLGKNAKLSIQELTWMKSYRDCAW 541
DB 588 LRSRSGVCVPAEARELREELIAKFLHMLMSYVVELLSFYYVETETQKRLFEYRKS 647
OY 542 LRSRSGVCVPAEARELREELIAKFLHMLMSYVVELLSFYYVETETQKRLFEYRKS 601
DB 648 VWSKQSIGIRHOLKRVOLRELSAEVROHREARPAALLTSRLRFIPKPDGLRPYVMDYV 707
OY 602 VWSKQSIGIRHOLKRVOLRELSAEVROHREARPAALLTSRLRFIPKPDGLRPYVMDYV 661
DB 708 VGARFRFRERKRELRSLRKALFSLVLYNRRARPGILGASVGLDIDHNAWTFVLRYVA 767
OY 662 VGARFRFRERKRELRSLRKALFSLVLYNRRARPGILGASVGLDIDHNAWTFVLRYVA 721
DB 768 QDPPELVKVDVYGAFTIPQDLTEVIAIIRKQNTCYCRYAVVQKAHGHYRKA 827
OY 722 QDPPELVKVDVYGAFTIPQDLTEVIAIIRKQNTCYCRYAVVQKAHGHYRKA 781
DB 828 KSHVSTLFDLQPYMROFVAHLQETSPLRDAVYIEQSSSLNEASSGLEDFVFLRPMCHAVR 887
OY 782 KSHVSTLFDLQPYMROFVAHLQETSPLRDAVYIEQSSSLNEASSGLEDFVFLRPMCHAVR 841
DB 888 IKGKSYVOCQIGIPQSIGISLTLCSLCYGMENKLPAGIRRDGLRLVDDFLVPHLTH 947
OY 842 IKGKSYVOCQIGIPQSIGISLTLCSLCYGMENKLPAGIRRDGLRLVDDFLVPHLTH 901
DB 948 AKFTLRTLVRGVPEXGVNLRKTYVNFVEDEALGTAFAVQMPAHGLPWCGLLDTXT 1007
OY 902 AKFTLRTLVRGVPEXGVNLRKTYVNFVEDEALGTAFAVQMPAHGLPWCGLLDTXT 961
DB 1008 LEVQDSYSSYARTSIRASLTFNRGFRAGNMRKLEGVLRKCHSLFLLDLOVNSIQTVCT 1067
OY 962 LEVQDSYSSYARTSIRASLTFNRGFRAGNMRKLEGVLRKCHSLFLLDLOVNSIQTVCT 1021
DB 1066 NIKYKILLQAYFFHACVLDLPFHQOYWKNPFTFLRYISTASLCSYLAKAKNAGSLGAK 1127
OY 1022 NIKYKILLQAYFFHACVLDLPFHQOYWKNPFTFLRYISTASLCSYLAKAKNAGSLGAK 1081
DB 1128 GAGGLPSEAVOMLCHQAFLLKTRHRVTVYVPLGSLTAQTQLSKRLGTTLTALEANA 1187
OY 1082 GAGGLPSEAVOMLCHQAFLLKTRHRVTVYVPLGSLTAQTQLSKRLGTTLTALEANA 1141
DB 1188 NPALPSDEFTIIL 1200
OY 1142 NPALPSDEFTIIL 1154

CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 600:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1285 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1285
CC OTHER INFORMATION: /note= "fusion protein composed of
CC OTHER INFORMATION: enterokinase cleavable, His tagged
CC OTHER INFORMATION: thiodoxin moiety and full length hTTF"
CC
SQ SEQUENCE 1285 AA; 143529 MW; 8449280 CN;
Query Match 99.8%; Score 8607; DB 18; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB 133 ASTORCVLLRTWEALAPATPAMPAPRCRAVNSLSRSHYREVLPLATFVRRLGPOGWRVLY 192
OY 2 ASGQCVLLRTWEALAPATPAMPAPRCRAVNSLSRSHYREVLPLATFVRRLGPOGWRVLY 61

Db 193 QRGDPAAERAIWAOCYVCPMDARPPAPSPRQVSCLEKELVARYLQRLCERGANVLA 252
Qy 62 QRGDPAAERAIWAOCYVCPMDARPPAPSPRQVSCLEKELVARYLQRLCERGANVLA 121
Db 253 GFALIDGARGGPEAFTTSVRSYLPNTVTDALRGSGAGLLLRVDDVYLHLLARCALF 312
Qy 122 GFALIDGARGGPEAFTTSVRSYLPNTVTDALRGSGAGLLLRVDDVYLHLLARCALF 181
Db 313 VLVAPSCAYOYCGPPLVYOLGAATOARPPHASGPRRRRLGCEBAMHVSREAGVPLGLPAP 372
Qy 182 VLVAPSCAYOYCGPPLVYOLGAATOARPPHASGPRRRRLGCEBAMHVSREAGVPLGLPAP 241
Db 373 GARRGGSASRSLPLPKRRPGRGAPEPERTPVQGSMAHPGRTGRSPRGCEVSPAPPA 432
Qy 242 GARRGGSASRSLPLPKRRPGRGAPEPERTPVQGSMAHPGRTGRSPRGCEVSPAPPA 301
Db 433 EEAATSLGALSGTRSHSPSVGRQHAGPPSTSRPPMDTFCPPVYATKHEFLYSSGDK 492
Qy 302 EEAATSLGALSGTRSHSPSVGRQHAGPPSTSRPPMDTFCPPVYATKHEFLYSSGDK 361
Db 493 QLRPSFLISSLRPSLTGARRLYETIFLGSRRPMGTPRRLPRLQRYOMRPLELLEGN 552
Qy 362 QLRPSFLISSLRPSLTGARRLYETIFLGSRRPMGTPRRLPRLQRYOMRPLELLEGN 421
Db 553 HAQCPYGLTKTHCPRLAAYTPAAGVCAREKPOGSVAPEEEDTPRRLVOLLRQHSPPM 612
Qy 422 HAQCPYGLTKTHCPRLAAYTPAAGVCAREKPOGSVAPEEEDTPRRLVOLLRQHSPPM 481
Db 613 QYGVFVRACLRLVPPGLMGRHNERRLRNTKFTISLGNKAKLSLOELJWMSYRDCAM 672
Qy 482 QYGVFVRACLRLVPPGLMGRHNERRLRNTKFTISLGNKAKLSLOELJWMSYRDCAM 541
Db 673 LRSRGVCVPAAEERLEETLAKELHMLASYVVELRSEFYVETEFQNRLEFFYKS 732
Qy 542 LRSRGVCVPAAEERLEETLAKELHMLASYVVELRSEFYVETEFQNRLEFFYKS 601
Db 733 VMSKLOSIGIRQHLKRVOLRELSAEVQRHREARPALITSRLREIPKPDGLRPIYNDYV 792
Qy 602 VMSKLOSIGIRQHLKRVOLRELSAEVQRHREARPALITSRLREIPKPDGLRPIYNDYV 661
Db 793 VGARTFRERKAERLTSVKALFSYLANTERARPPGLGASVGLDDIRAMWTFYLAVRA 852
Qy 662 VGARTFRERKAERLTSVKALFSYLANTERARPPGLGASVGLDDIRAMWTFYLAVRA 721
Db 853 QDPPELYEFVVDVGTAVDTIPQDLTEVIAIIRKQNTYCVRAYVQKAAHGVRAAF 912
Qy 722 QDPPELYEFVVDVGTAVDTIPQDLTEVIAIIRKQNTYCVRAYVQKAAHGVRAAF 781
Db 913 KSHVSTLTDLPYMRQFAVHAHQETSPLRDVAVIEOSSSLNEASSGLFDFVFLRFMCNHA 972
Qy 782 KSHVSTLTDLPYMRQFAVHAHQETSPLRDVAVIEOSSSLNEASSGLFDFVFLRFMCNHA 841
Db 973 IRGKSIYVOCGIGIPGOSIILSTLCSLCYGDMEKNKLEFAGIRRDGLLRVLDVDFLLVPHLTH 1032
Qy 842 IRGKSIYVOCGIGIPGOSIILSTLCSLCYGDMEKNKLEFAGIRRDGLLRVLDVDFLLVPHLTH 901
Db 1033 AKTELRTLVRCVPEVCGVNVNLRKTYVNVPEVDEALGATFQVOMPAHGLFPPCGGLLDRT 1092
Qy 902 AKTELRTLVRCVPEVCGVNVNLRKTYVNVPEVDEALGATFQVOMPAHGLFPPCGGLLDRT 961
Db 1093 LEVOSDYSSYARTSIRASLTFRNGFKAGRNARRKLFGLVLRKCHSLFLDLOVNSIQTYCT 1152
Qy 962 LEVOSDYSSYARTSIRASLTFRNGFKAGRNARRKLFGLVLRKCHSLFLDLOVNSIQTYCT 1021
Db 1153 NIYKILLQAVRFHACVQLQEPFHQVWKNPTEFLRVISDTASLCYSILKANAGNSLGAK 1212
Qy 1022 NIYKILLQAVRFHACVQLQEPFHQVWKNPTEFLRVISDTASLCYSILKANAGNSLGAK 1081
Db 1213 GAAPPLPSEAVQMLCHQAFLLKTRHRTTYPLGLSLTAOTQUSRKLPGLTTLALEMAA 1272
Qy 1082 GAAPPLPSEAVQMLCHQAFLLKTRHRTTYPLGLSLTAOTQUSRKLPGLTTLALEMAA 1141
Db 1273 NPALPDSDFKTLID 1285

Qy 1142 NPALPDSDFKTLID 1154
RESULT 9
ID US-08-911-312-32 STANDARD: PRT: 1285 AA.
AC xxxxxx
XX
XX
XX
DE Sequence 32, Application US/08911312
XX
XX Sequence 32, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911,312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-002500US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 32:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1285 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:

CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 1285 AA; 1435280 MW; 8449280 CN;
Query Match 99.8%; Score 8607; DB 13; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 133 ASTORCVLLRTWEALAPATPMPAPRCRAVNSILRSHRYETLPLATVYRRLGPGCMGLY 192
QY 2 ASGGCVLLRTWEALAPATPMPAPRCRAVNSILRSHRYETLPLATVYRRLGPGCMGLY 61
Db 193 QRGDPAFRAVLAOCVCPMDAPPPAPSPROYSCLEKELVARTLQRLCEGANVIAF 252
QY 62 QRGDPAFRAVLAOCVCPMDAPPPAPSPROYSCLEKELVARTLQRLCEGANVIAF 121
Db 253 GFALLDGAARGGPEAFPTSVRSYLPNTVTDALRGSGAGLLRRYGDVYLHLLARCALE 312
QY 122 GFALLDGAARGGPEAFPTSVRSYLPNTVTDALRGSGAGLLRRYGDVYLHLLARCALE 181
Db 313 VLVAPECAIYOCGPPYLQGAATQARPPPHASGPRRLGCEPAMNHSYREAGVPLGLAP 372
QY 182 VLVAPECAIYOCGPPYLQGAATQARPPPHASGPRRLGCEPAMNHSYREAGVPLGLAP 241
Db 373 GARRRGSGASRSLPLKPRPRGAPEPERTPVGSGSMAHPGRTGSDRGFCVSPAPPA 432
QY 242 GARRRGSGASRSLPLKPRPRGAPEPERTPVGSGSMAHPGRTGSDRGFCVSPAPPA 301
Db 433 BEATLEGALSTGRSHSPVSGRQHAGPPTSRPPRMDPCPPYAEIKHPLSSGDK 492
QY 302 BEATLEGALSTGRSHSPVSGRQHAGPPTSRPPRMDPCPPYAEIKHPLSSGDK 361
Db 493 QLRPSFLSLSPSLTGARLVEYTFGLSRPMPGTPRRPLPQRYQMRLPFLLELGN 552
QY 362 QLRPSFLSLSPSLTGARLVEYTFGLSRPMPGTPRRPLPQRYQMRLPFLLELGN 421
Db 553 HAOCYGVLLKHCPLRAVPAVPAAGVCAKEKPGSVAAPEEDTDPRLYOLLRHSSPW 612
QY 422 HAOCYGVLLKHCPLRAVPAVPAAGVCAKEKPGSVAAPEEDTDPRLYOLLRHSSPW 481
Db 613 QYVGVRACLRRLVPGMLSGSRHNERFLNRTKFTISLGHAKLSLOELTWMMSYRDCAM 672
QY 482 QYVGVRACLRRLVPGMLSGSRHNERFLNRTKFTISLGHAKLSLOELTWMMSYRDCAM 541
Db 673 LRSRPGVCVPAEHRLEBEILAKFLHMLSVYVVELLSEFYVETTFQKNRLEFFYRKS 732
QY 542 LRSRPGVCVPAEHRLEBEILAKFLHMLSVYVVELLSEFYVETTFQKNRLEFFYRKS 601
Db 733 VMSKLOSIGIRQHLKRVOLRELSAEVQRHREARPALTLSTRLETPKPGGLRPYNNMIV 792
QY 602 VMSKLOSIGIRQHLKRVOLRELSAEVQRHREARPALTLSTRLETPKPGGLRPYNNMIV 661
Db 793 VGARTFRREKRAERLTSRKALFSVLNTERARPPGLGASVGLDDIHRAMRTFLYARA 852
QY 662 VGARTFRREKRAERLTSRKALFSVLNTERARPPGLGASVGLDDIHRAMRTFLYARA 721
Db 853 QDPPELYFVKVDTGAYDTIPQDRLETVIASIIRPQNTYCVRRYAVVQKAAGHYRAAF 912
QY 722 QDPPELYFVKVDTGAYDTIPQDRLETVIASIIRPQNTYCVRRYAVVQKAAGHYRAAF 781
Db 913 KSHVSTLIDLOPYMROFVAHLOETSPLRVAVYIEOSSSLNENSSGGLFDFLTFMKHHAAR 972
QY 782 KSHVSTLIDLOPYMROFVAHLOETSPLRVAVYIEOSSSLNENSSGGLFDFLTFMKHHAAR 841
Db 973 IRGKSYVOCQIPQGSILSTLCSLCYGMENKLFAGIRDDILLRLYDFFLTPHLLH 1032
QY 842 IRGKSYVOCQIPQGSILSTLCSLCYGMENKLFAGIRDDILLRLYDFFLTPHLLH 901
Db 1033 AKTFLRTLVRGVPEYGCVVNLKRTYVNFVEDEALGTAFAVOMPAHGLEPPMCGLLDTRT 1092
QY 902 AKTFLRTLVRGVPEYGCVVNLKRTYVNFVEDEALGTAFAVOMPAHGLEPPMCGLLDTRT 961
Db 1093 LEVQSDSYARTSIRASLTFRNGFAGNMNRKLFGLVRLKCHSLFLDLOVNSLQTYCT 1152

QY 962 LEVQSDSYARTSIRASLTFRNGFAGNMNRKLFGLVRLKCHSLFLDLOVNSLQTYCT 1021
Db 1153 NITYKILLQATRFNACVQLQLPFHQOVWKNPFELRVISDFASLCTYSILKANAAGMSIGAK 1212
QY 1022 NITYKILLQATRFNACVQLQLPFHQOVWKNPFELRVISDFASLCTYSILKANAAGMSIGAK 1081
Db 1213 GAAGPLSEAVQWICHQAFELKLRHRYTVPLGLSRTAQTQLSRKLPQTTLALEMAA 1272
QY 1082 GAAGPLSEAVQWICHQAFELKLRHRYTVPLGLSRTAQTQLSRKLPQTTLALEMAA 1141
Db 1273 NPALPSDEFTILD 1285
QY 1142 NPALPSDEFTILD 1154
RESULT 10
ID US-08-912-951-314 STANDARD; PRT; 1285 AA.
AC xxxxxx
XX
XX
DE Sequence 314, Application US/08912951
CC
CC Sequence 314, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morlin, Gregg B.
CC APPLICANT: Harley, Calvin H.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002600US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 314:
SEQUENCE CHARACTERISTICS:
LENGTH: 1285 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 1285 AA: 143529 MW: 8449280 CN:
Query Match 99.8%; Score 8607; DB 13; Length 1285;
Best Local Similarity 99.9%; Pred. No. 0.00e+00;
Matches 1152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 133 ASTORCVLLFTWEALPATPAMPAPRCRAVRSILRSHYREVLPLATVRRLGPGQWRLV 192
QY 2 ASGRCVLLFTWEALPATPAMPAPRCRAVRSILRSHYREVLPLATVRRLGPGQWRLV 61
Db 193 QRCDAFALVAOCIVCPMDARPPAPSPFQVSCLELYARVQLRCERGAKNVLA 252
QY 62 QRCDAFALVAOCIVCPMDARPPAPSPFQVSCLELYARVQLRCERGAKNVLA 121
Db 253 GFALIDGARGGPEAFTTSVRSYLPNTVTDALRGSGANGLLRRVGGDVLVHLARCALF 312
QY 122 GFALIDGARGGPEAFTTSVRSYLPNTVTDALRGSGANGLLRRVGGDVLVHLARCALF 181
Db 313 VLVAWSCAYVCYGPPLVQLGATQARPPHASGPRRLGECERAMNHSVREGVPLGLPAP 372
QY 182 VLVAWSCAYVCYGPPLVQLGATQARPPHASGPRRLGECERAMNHSVREGVPLGLPAP 241
Db 373 GARRRGGSASRLPLPKPRRGAAPEPERTVGGGSMHAPGTRGSPRGCVVSPARPA 432
QY 242 GARRRGGSASRLPLPKPRRGAAPEPERTVGGGSMHAPGTRGSPRGCVVSPARPA 301
Db 433 EEAATSLGALSSTGTHSHSVGRQHAGPSTSRPPRPMDTCPVYATKHFLLYSSGDK 492
QY 302 EEAATSLGALSSTGTHSHSVGRQHAGPSTSRPPRPMDTCPVYATKHFLLYSSGDK 361
Db 493 QLRSPFLSLSRPLTGARLVETIFLGRPMWGTPTRLRLQRYQMPLELLELGN 552
QY 362 QLRSPFLSLSRPLTGARLVETIFLGRPMWGTPTRLRLQRYQMPLELLELGN 421
Db 553 HAQCPYGLTKHCPPLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 612
QY 422 HAQCPYGLTKHCPPLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 481
Db 613 QVYGFVRACLRRLVPGILMGRHNERFLRNTKFFISLGKHAKLSDLELTWKMVSVDCAW 672
QY 482 QVYGFVRACLRRLVPGILMGRHNERFLRNTKFFISLGKHAKLSDLELTWKMVSVDCAW 541
Db 673 LRRSPGVGCVAHAERLRBEELIAFLHMLSVYVELLSRFYTTETTFQNNRLEFFTKS 732
QY 542 LRRSPGVGCVAHAERLRBEELIAFLHMLSVYVELLSRFYTTETTFQNNRLEFFTKS 601
Db 733 VMSLQSLGIGIHOHLKRVOLRELSEAEVQRHREARPAALTSRLRFPKPDGRPIVNMNVY 792
QY 602 VMSLQSLGIGIHOHLKRVOLRELSEAEVQRHREARPAALTSRLRFPKPDGRPIVNMNVY 661
Db 793 VGATTFRRERKRAERLTSVKALFSVLNTERARBPGLGASVYLGLDJIHRAWTFVLRYRA 852
QY 662 VGATTFRRERKRAERLTSVKALFSVLNTERARBPGLGASVYLGLDJIHRAWTFVLRYRA 721
Db 853 QDPPELYFYVVDVYGAADTIPQDRLETVASIIKPQNTYVYKRAYVQKAAHGHVRAAF 912
QY 722 QDPPELYFYVVDVYGAADTIPQDRLETVASIIKPQNTYVYKRAYVQKAAHGHVRAAF 781
Db 913 KSHVSTLTDLPYVRQFVAHQIETSPRLDAVAVIQQSSSLNEASSGLDFDVLRFMCHNAVR 972

|||||
QY 782 KSHVSTLTDLPYVRQFVAHQIETSPRLDAVAVIQQSSSLNEASSGLDFDVLRFMCHNAVR 841
Db 973 IRGSIYVQCGIPIQGSILSTLCSLCYGDMEKNKIFAGIRRDGLLRVDDFLVTPHLTH 1032
QY 842 IRGSIYVQCGIPIQGSILSTLCSLCYGDMEKNKIFAGIRRDGLLRVDDFLVTPHLTH 901
Db 1033 AKTFPLTVGPEYGVNLRKTVNFPVEDEALGSTAVQMPAHGLFPMCGLLDTRT 1092
QY 902 AKTFPLTVGPEYGVNLRKTVNFPVEDEALGSTAVQMPAHGLFPMCGLLDTRT 961
Db 1093 LEVQSDYSYARTSIRASLTFFNRFKAGRMRRKLEGVRLKCHSLFLDLQVNSLQTVCT 1152
QY 962 LEVQSDYSYARTSIRASLTFFNRFKAGRMRRKLEGVRLKCHSLFLDLQVNSLQTVCT 1021
Db 1153 NIYKILLQAVRFHACVQLPFIHQVKNKPTFLRVISDTASISLCYSLKAKNAGMSIGAK 1212
QY 1022 NIYKILLQAVRFHACVQLPFIHQVKNKPTFLRVISDTASISLCYSLKAKNAGMSIGAK 1081
Db 1213 GAAGPLPSEAVQWICHQAFLLKLRHRTVYVPLGLSRTAOTOLSRKLPQTTLTALEAA 1272
QY 1082 GAAGPLPSEAVQWICHQAFLLKLRHRTVYVPLGLSRTAOTOLSRKLPQTTLTALEAA 1141
Db 1273 NPALPSDKTILD 1285
QY 1142 NPALPSDKTILD 1154
RESULT 11
ID US-08-911-312-55 STANDARD; PRT: 1407 AA.
AC xxxxxx
XX
XX
DE Sequence 55, Application US/08911312
XX
CC Sequence 55, Application US/08911312
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Langner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Hatley, Calvin B.
CC APPLICANT: Andrews, William
CC TITLE OF INVENTION: Telomerase Reverse Transcriptase
CC NUMBER OF SEQUENCES: 170
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/911,312
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Einhorn, Gregory P.
CC REGISTRATION NUMBER: 38,440
CC REFERENCE/DOCKET NUMBER: 015389-00250005
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 55:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1407 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1407 AA; 157668 MW; 10134798 CN;

Query Match 99.6%; Score 8590; DB 13; Length 1407;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLRTWELALPAIPAPAPRAVCRAVNSLSHREVLPLATFYRRLLPGCGRWLV 314
|||
|||
QY 2 ASGORCVLLRTWELALPAIPAPAPRAVCRAVNSLSHREVLPLATFYRRLLPGCGRWLV 61
|||
|||
Db 315 ORGDDPAFPAALVAOCVCPMDAPPPAPPSFQVSCLELVARVLCRCERGAKNVLA 374
|||
|||
QY 62 ORGDDPAFPAALVAOCVCPMDAPPPAPPSFQVSCLELVARVLCRCERGAKNVLA 121
|||
|||
Db 375 GFALLDGAAGGPEAFTTSVSYLPTVTVDALRGSGAMGILLRVGDDVYLHLLARCA 434
|||
|||
QY 122 GFALLDGAAGGPEAFTTSVSYLPTVTVDALRGSGAMGILLRVGDDVYLHLLARCA 181
|||
|||
Db 435 VVAPSCAVQVCGPPLYQLGAATQAPPPHAGSPRRRLCCERAMNSVREAGVPLGLPAP 494
|||
|||
QY 182 VVAPSCAVQVCGPPLYQLGAATQAPPPHAGSPRRRLCCERAMNSVREAGVPLGLPAP 241
|||
|||
Db 495 GARRRGSGASRSPLPKRRRGGAAPERPTVQGSMAHGRTRGSDGFCVSPARPA 554
|||
|||
QY 242 GARRRGSGASRSPLPKRRRGGAAPERPTVQGSMAHGRTRGSDGFCVSPARPA 301
|||
|||
Db 555 EEAATSEGAALSGTRASHHPVSGROHAGPPSTSRPPMDTCPBPVVAETKHFLYSGDKE 614
|||
|||
QY 302 EEAATSEGAALSGTRASHHPVSGROHAGPPSTSRPPMDTCPBPVVAETKHFLYSGDKE 361
|||
|||
Db 615 QLRPSFLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLLPRLPORYWOMRPLLELGN 674
|||
|||
QY 362 QLRPSFLSSLRPSLTGARRLVETIFLGSRPWMPGTPRRLLPRLPORYWOMRPLLELGN 421
|||
|||
Db 675 HAOCPPVLLKTHCPRLAANTPRAAGCAREKPGGSVAABEEEDTDRRLVOLLROHSSPW 734
|||
|||
QY 422 HAOCPPVLLKTHCPRLAANTPRAAGCAREKPGGSVAABEEEDTDRRLVOLLROHSSPW 481
|||
|||
Db 735 QYVGFYRACLRLRVPGLGSRHNERFLRNTKFTSLGKHAFLSQDELTKWMSVDCAM 794
|||
|||
QY 482 QYVGFYRACLRLRVPGLGSRHNERFLRNTKFTSLGKHAFLSQDELTKWMSVDCAM 541
|||
|||
Db 795 LRRSPGVGVPAAEHRLREELIAKFLHMLMSYVVELLSFFVYETTFQKNLFFYRPS 854
|||
|||
QY 542 LRRSPGVGVPAAEHRLREELIAKFLHMLMSYVVELLSFFVYETTFQKNLFFYRPS 601
|||
|||
Db 855 VMSKLSISIRQHLKRVQLRELSEALVROHREARPAALLISRLRFIPKPGCLRIYVMDYV 914
|||
|||

QY 602 VMSKLSISIRQHLKRVQLRELSEALVROHREARPAALLISRLRFIPKPGCLRIYVMDYV 661
|||
|||
Db 915 VQARTFRREKRAERLTSRVKALFSVLNTERARPPGLIGASVLGLDDIHRAMRTFYLVRBA 974
|||
|||
QY 662 VQARTFRREKRAERLTSRVKALFSVLNTERARPPGLIGASVLGLDDIHRAMRTFYLVRBA 721
|||
|||
Db 975 ODPPELVYKVDYNGAVYTIPODRLTEVIASIIKQNTYCVRRYAVQKAHGHVRAKAF 1034
|||
|||
QY 722 ODPPELVYKVDYNGAVYTIPODRLTEVIASIIKQNTYCVRRYAVQKAHGHVRAKAF 781
|||
|||
Db 1035 KSHVSTFDLPYMRQFVNLQETSPLRDAVYIEOSSLNKASGLFDFLRFMCHAAVR 1094
|||
|||
QY 782 KSHVSTFDLPYMRQFVNLQETSPLRDAVYIEOSSLNKASGLFDFLRFMCHAAVR 841
|||
|||
QY 1095 INKSYVOCGIPQGSILSTLCSLCYGDMEKLFAGIRRDGILLRLVDFLLVTPHLTH 1154
|||
|||
Db 842 INKSYVOCGIPQGSILSTLCSLCYGDMEKLFAGIRRDGILLRLVDFLLVTPHLTH 901
|||
|||
QY 1155 ARTFLRTLVRGVPEYGCYVNLKRTYVNFVEDEALGTAFFVQMPAHGLPPMGILLDTPT 1214
|||
|||
Db 902 ARTFLRTLVRGVPEYGCYVNLKRTYVNFVEDEALGTAFFVQMPAHGLPPMGILLDTPT 961
|||
|||
Db 1215 LEVQSDYSYARTSIRASVTENRGFAGRNMRKLFGLRLCHSLFLDLQVNSLOTVCT 1274
|||
|||
QY 962 LEVQSDYSYARTSIRASVTENRGFAGRNMRKLFGLRLCHSLFLDLQVNSLOTVCT 1021
|||
|||
Db 1275 NIYKILLQAYRFAVCVLDLPFHQOYWKNPTEFLRYISDTASLCYSILAKNAGSLGAK 1334
|||
|||
QY 1022 NIYKILLQAYRFAVCVLDLPFHQOYWKNPTEFLRYISDTASLCYSILAKNAGSLGAK 1081
|||
|||
Db 1335 GAAGPLPSAIVOMLCHQALFLKTRHRTYVPLLSGLRTAQOLSKLPGTTLTALCAA 1394
|||
|||
QY 1082 GAAGPLPSAIVOMLCHQALFLKTRHRTYVPLLSGLRTAQOLSKLPGTTLTALCAA 1141
|||
|||
Db 1395 NPALPSDFKTIID 1407
|||
|||
QY 1142 NPALPSDFKTIID 1154
|||
|||

RESULT 12
ID US-08-974-549-628 STANDARD: PRT: 1407 AA.
AC xxxxxx
XX
XX
XX
XX
DE Sequence 628, Application US/08974549
XX
XX
CC Sequence 628, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morley, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885
CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 628:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1407 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: Protein
CC LOCATION: 1..1407
CC OTHER INFORMATION: /note="fusion protein composed of
CC OTHER INFORMATION: enhanced green fluorescent protein
CC OTHER INFORMATION:
CC OTHER INFORMATION: 5' untranslated region of hTERT mRNA and
CC OTHER INFORMATION: hTERT protein sequence"
CC SEQUENCE 1407 AA; 157668 MM; 10134798 CN;
Query Match 99.6%; Score 8590; Db 18; Length 1407;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLRTWEALAPAPAMPAPRCRAVSLRSHYREVLPLATFVRRLGPGGWRIV 314
QY 2 ASGQCVLLRTWEALAPAPAMPAPRCRAVSLRSHYREVLPLATFVRRLGPGGWRIV 61
Db 315 QRGDPAFRLVAQCIVCPMDARPPAPSPFQVSCLELVARVQLRCEGAKNVIAF 374
QY 62 QRGDPAFRLVAQCIVCPMDARPPAPSPFQVSCLELVARVQLRCEGAKNVIAF 121
Db 375 GFALLDGAAGGPEAFFTTSVRSYLPNTVTDALRGSGAMGLLRRVGDDVLVHLARCALE 434
QY 122 GFALLDGAAGGPEAFFTTSVRSYLPNTVTDALRGSGAMGLLRRVGDDVLVHLARCALE 181

Db 435 VLVAISCAYQVCGPPLVYOLGATQAPPPHASGPRRLGECERAMNHSVEAGVPLGLPAP 494
QY 192 VLVAISCAYQVCGPPLVYOLGATQAPPPHASGPRRLGECERAMNHSVEAGVPLGLPAP 241
Db 495 GARRRGSSASNSLPLPKRRRGAAPEPERTVGGQSWAHPGRTGPSDRGFCVSPAPPA 554
QY 242 GARRRGSSASNSLPLPKRRRGAAPEPERTVGGQSWAHPGRTGPSDRGFCVSPAPPA 301
Db 555 BEATLEGALSGTRSHSPVGRHAGPSTSRPPMDTCCPPYATKTFVSSGDK 614
QY 302 BEATLEGALSGTRSHSPVGRHAGPSTSRPPMDTCCPPYATKTFVSSGDK 361
Db 615 QLRPSFLSLRPSLTGARRLVETIFLGRPMWGTPTPLRLPORVQMPPLFELIGN 674
QY 362 QLRPSFLSLRPSLTGARRLVETIFLGRPMWGTPTPLRLPORVQMPPLFELIGN 421
Db 675 HAQCYGVLLTCHPCLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 734
QY 422 HAQCYGVLLTCHPCLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 481
Db 735 QYGVVRACLRLVPPGILMGSRHNRRLRMTKKFISLGRKAKLSLOELTKMMSVROCAW 794
QY 482 QYGVVRACLRLVPPGILMGSRHNRRLRMTKKFISLGRKAKLSLOELTKMMSVROCAW 541
Db 795 LRSBGVCVPAAEHRLKEELIAFLHMLMSVYVELLSFEYVETTFQKNRLFYPPS 854
QY 542 LRSBGVCVPAAEHRLKEELIAFLHMLMSVYVELLSFEYVETTFQKNRLFYPPS 601
Db 855 VWSKLSIGIRHQLKRVOLRELSEAEVQRHREARPAITLSRLRPIKPDGLPIVMDYV 914
QY 602 VWSKLSIGIRHQLKRVOLRELSEAEVQRHREARPAITLSRLRPIKPDGLPIVMDYV 661
Db 915 VGARFREREKRAERLTSVKALFSVLANERARRPGLIGASVIGDDIHARARTFVLRRRA 974
QY 662 VGARFREREKRAERLTSVKALFSVLANERARRPGLIGASVIGDDIHARARTFVLRRRA 721
Db 975 QDPPELTFVAVDYGADTIPODRLEVIASIIKPONTYCVRRAYVQKAHGHVRAAF 1034
QY 722 QDPPELTFVAVDYGADTIPODRLEVIASIIKPONTYCVRRAYVQKAHGHVRAAF 781
Db 1035 KSHVSTLDLPYMRQVFAHLQETSPLRDAVYIEOSSLNAASSGLDFVFLRPMCHNAVR 1094
QY 782 KSHVSTLDLPYMRQVFAHLQETSPLRDAVYIEOSSLNAASSGLDFVFLRPMCHNAVR 841
Db 1095 IRGKSYVOCQGIPOGSIILSTLCSLCYGDMDENKLFAGIRBDGLLRLVDDDLVTPHLTH 1154
QY 842 IRGKSYVOCQGIPOGSIILSTLCSLCYGDMDENKLFAGIRBDGLLRLVDDDLVTPHLTH 901
Db 1155 AKTEFLRTLVRGVPEYGCYVNLRTVYVNFVEDEALGTAFFQMPAHGLFPACGLLDTRT 1214
QY 902 AKTEFLRTLVRGVPEYGCYVNLRTVYVNFVEDEALGTAFFQMPAHGLFPACGLLDTRT 961
Db 1215 LEVQSDYSYVARTSIRASVTENRNGFKAGRNMRKILFVYLKCHSLFDLDVNSLQVCT 1274
QY 962 LEVQSDYSYVARTSIRASVTENRNGFKAGRNMRKILFVYLKCHSLFDLDVNSLQVCT 1021
Db 1275 NIYKTLILQAFRAHACVQLQFFHQVWKNPTEFLRVISDTASISLCTIKANAGSGAK 1334
QY 1022 NIYKTLILQAFRAHACVQLQFFHQVWKNPTEFLRVISDTASISLCTIKANAGSGAK 1081
Db 1335 GAAGPLPSEAVOMLCHQAFLLKTRHRTVYVPLLSLRTAQTOLSRKLPGTTILALEAA 1394
QY 1082 GAAGPLPSEAVOMLCHQAFLLKTRHRTVYVPLLSLRTAQTOLSRKLPGTTILALEAA 1141
Db 1395 NPALPDSDEKTIID 1407
QY 1142 NPALPDSDEKTIID 1154

RESULT 13
ID US-08-912-951-334 STANDARD; PRT: 1407 AA.
XX
AC xxxxxx

XX XX Sequence 334, Application US/08912951
XX DE
XX CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-00260005
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 334:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1407 amino acids
CC TYPE: amino acid
CC STRANDEDNESS:
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1407 AA; 157668 MW; 10134798 CN;

Query Match 99.6%; Score 8590; DB 13; Length 1407;
Best Local Similarity 99.7%; Pred. No. 0.00e+00;
Matches 1150; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Db 255 ASTORCVLLRTWEALAPATPAMPAPRCRAVRSLSRSHREVLPATFVRRLGPGWRLV 314
Qy 2 ASGORCVLLRTWEALAPATPAMPAPRCRAVRSLSRSHREVLPATFVRRLGPGWRLV 61
Db 315 QRGDPAAFRALVAOCLVCPWDARPPPAASFROYSCLELVARYLQLCERGAKNVIAF 374
Qy 62 QRGDPAAFRALVAOCLVCPWDARPPPAASFROYSCLELVARYLQLCERGAKNVIAF 121
Db 375 GFALLDARGGPPPEAFETTSVRSYLPNTYTDALRSGAGLLRRGGDVLVHLARCLF 434
Qy 122 GFALLDARGGPPPEAFETTSVRSYLPNTYTDALRSGAGLLRRGGDVLVHLARCLF 181
Db 435 VLVAPSCAYOCGPPPLYOLGAATQARPPPHASGPRRLGCERAMNHSVREAGVPLGLPAP 494
Qy 132 VLVAPSCAYOCGPPPLYOLGAATQARPPPHASGPRRLGCERAMNHSVREAGVPLGLPAP 241
Db 495 GARRRGASASRLPLKPRRGAAPERTPTVGGQSWAHPCRTGSPDRGCVSPAPPA 554
Qy 242 GARRRGASASRLPLKPRRGAAPERTPTVGGQSWAHPCRTGSPDRGCVSPAPPA 301
Db 555 BEATSLGALSGTRSHSPSVGRHAGPSTSRPPRMDTCCPYATATKHFVLSGDK 614
Qy 302 BEATSLGALSGTRSHSPSVGRHAGPSTSRPPRMDTCCPYATATKHFVLSGDK 361
Db 615 QLRPFLLSLRPLTGARLVETFTLGRPMMPETPRRLPDRYQOMPLFELIGN 674
Qy 362 QLRPFLLSLRPLTGARLVETFTLGRPMMPETPRRLPDRYQOMPLFELIGN 421
Db 675 HAQCPYGLTKHCPRLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 734
Qy 422 HAQCPYGLTKHCPRLRAAVTPAAGVCAREKPGQSVAAPEEDTDPRRLVOLLROHSSPW 481
Db 735 QYGGFVRACTRLVPPGMLGSRHNERRLRNTKRTISLGAKAKLSQLTMTKMSVRDCAW 794
Qy 482 QYGGFVRACTRLVPPGMLGSRHNERRLRNTKRTISLGAKAKLSQLTMTKMSVRDCAW 541
Db 795 LRRSPGVCVPAAEHRLREELIAKFLHMLSYVVELRSFFVYETETFOKNRLFFYRPS 854
Qy 542 LRRSPGVCVPAAEHRLREELIAKFLHMLSYVVELRSFFVYETETFOKNRLFFYRPS 601
Db 855 VWSKLQSIGIRQHLKRVOLRELSEAEVQRHREARALLTSRLRFLPKPDGLRPIVMDYV 914
Qy 602 VWSKLQSIGIRQHLKRVOLRELSEAEVQRHREARALLTSRLRFLPKPDGLRPIVMDYV 661
Db 915 VGARTFRREKRAERLTSVKALFSVLNTERARPCGLCAVLGLDDIRARMTFVLRA 974
Qy 662 VGARTFRREKRAERLTSVKALFSVLNTERARPCGLCAVLGLDDIRARMTFVLRA 721
Db 975 QDPPELTFVAVDTGAVDTIPQDLTEVIAIIRPQNTYCVRAVYVOKAHGVRKAF 1034
Qy 722 QDPPELTFVAVDTGAVDTIPQDLTEVIAIIRPQNTYCVRAVYVOKAHGVRKAF 781
Db 1035 KSHVSTLDLPYMRFOFAHLQETSPLRDAAVIEQSSSLNEASSGLFVFLRFMCHHAVR 1094
Qy 782 KSHVSTLDLPYMRFOFAHLQETSPLRDAAVIEQSSSLNEASSGLFVFLRFMCHHAVR 841
Db 1095 IRGKSYYOCGIPQGSILSTLCSLCYDMDENKLFAGIRQGLRLRLVYDDELVTPLHLTH 1154
Qy 842 IRGKSYYOCGIPQGSILSTLCSLCYDMDENKLFAGIRQGLRLRLVYDDELVTPLHLTH 901
Db 1155 AKTEFLRTLVKGVPEYGVVNLKTYVNFVDEALGTAFAVOMPAHGLFPMCGLLDRT 1214
Qy 902 AKTEFLRTLVKGVPEYGVVNLKTYVNFVDEALGTAFAVOMPAHGLFPMCGLLDRT 961
Db 1215 LEVQSDYSYARTSIRASVTENRFGKAGRNMRKLFGLVRLKCHSLFLDLQVNSIQTYCT 1274
Qy 962 LEVQSDYSYARTSIRASVTENRFGKAGRNMRKLFGLVRLKCHSLFLDLQVNSIQTYCT 1021
Db 1275 NIYKILLIQATRFACVATQLPPEHQVWKNPTFFLAVISDTSLSCTIILKANNAGSLAK 1334
Qy 1022 NIYKILLIQATRFACVATQLPPEHQVWKNPTFFLAVISDTSLSCTIILKANNAGSLAK 1081
Db 1335 GAAGPLPSEAVQWLOCHQAFLLKLTGRHRTVYVPLIGSLTATQOLSRKLPGLTTLAEAA 1394

|||||
QY 1082 GAAPLSEAVQWCHQAFILKTRHRTYTPLLGSLTFAQTQSRKLPGLTALAEAA 1141
Db 1395 NPALPSDEKTLTD 1407
QY 1142 NPALPSDEKTLTD 1154
RESULT 14
ID US-08-974-549-2 STANDARD: PRT: 1132 AA.
AC xxxxxx
XX
XX
XX
DE Sequence 2, Application US/08974549
XX
CC Sequence 2, Application US/08974549
CC GENERAL INFORMATION:
CC APPLICANT: Cech, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin B.
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: Human Telomerase Catalytic Subunit
CC NUMBER OF SEQUENCES: 726
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, Eighth Floor
CC City: San Francisco
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94111-3834
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/974,549
CC FILING DATE: 19-NOV-1997
CC CLASSIFICATION: 536
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/911,312
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/912,951
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/915,503
CC FILING DATE: 14-AUG-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17618
CC FILING DATE: 01-OCT-1997
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: WO PCT/US97/17885

CC FILING DATE: 01-OCT-1997
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph Ted
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002610US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1132 AA, 126995 MW, 6588209 CN;
SQ
Query Match 98.28; Score 8465; DB 18; Length 1132;
Best Local Similarity 100.08; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MPRAPRCRAVRSLSLRSHRYREVLPLATFVRRLGPGQWRVLVGRGPPAARALVAQCIVCPW 60
QY 23 MPRAPRCRAVRSLSLRSHRYREVLPLATFVRRLGPGQWRVLVGRGPPAARALVAQCIVCPW 82
Db 61 DARPPAPSPFROYSCKEIVARYLQRLCERGAKNVLAFGFALDGAAGGPPFAFTTSVR 120
QY 83 DARPPAPSPFROYSCKEIVARYLQRLCERGAKNVLAFGFALDGAAGGPPFAFTTSVR 142
Db 121 SYLPNTVTDALRGSGAGLLRLRYGDDVLVHLARCALFVLVAPSCAYQVCGPPLYQLGA 180
QY 143 SYLPNTVTDALRGSGAGLLRLRYGDDVLVHLARCALFVLVAPSCAYQVCGPPLYQLGA 202
Db 181 ATQARPPPHASGPRRLRCERANNSYREAGVPLGLPAPARRGSGASSLPLPKPRR 240
QY 203 ATQARPPPHASGPRRLRCERANNSYREAGVPLGLPAPARRGSGASSLPLPKPRR 262
Db 241 GAAPERTPVGQSNAPGRTGSPDRCFCVVSAPAPAEATSLGALGTRSHSVG 300
QY 263 GAAPERTPVGQSNAPGRTGSPDRCFCVVSAPAPAEATSLGALGTRSHSVG 322
Db 301 RQHHAGPSTSRPPMDTDCPPYATKHFLLYSSGDEQLRPSFLLSLRPSLTGARL 360
QY 323 RQHHAGPSTSRPPMDTDCPPYATKHFLLYSSGDEQLRPSFLLSLRPSLTGARL 382
Db 361 VETFFLSRPMGTPRRLRPLPORTQWMPRLLELGNNAOCYGVLLTTCPELRAVT 420
QY 383 VETFFLSRPMGTPRRLRPLPORTQWMPRLLELGNNAOCYGVLLTTCPELRAVT 442
Db 421 PAAGVCAREKPOGSVAAPDEEDTPRLVOLLROHSSPMQVGVFVACLRRLVPPGLMG 480
QY 443 PAAGVCAREKPOGSVAAPDEEDTPRLVOLLROHSSPMQVGVFVACLRRLVPPGLMG 502
Db 481 RHNERRLRNTKKEISLGKHAKLISLOELTWKMSYRDCAMLRSPGVGCVAAEHRLREEI 540
QY 503 RHNERRLRNTKKEISLGKHAKLISLOELTWKMSYRDCAMLRSPGVGCVAAEHRLREEI 562
Db 541 LAKFLHMLSYVYVELLRFFYTTETTFQKNRLEFFYKSWSKIOSIGIIOHLKRVOLRE 600
QY 563 LAKFLHMLSYVYVELLRFFYTTETTFQKNRLEFFYKSWSKIOSIGIIOHLKRVOLRE 622
Db 601 LSEAEVQOHEARPALTSRLRFIPKPDGLRPIYNMUYVGARFREREKRAERLTSVKA 660
QY 623 LSEAEVQOHEARPALTSRLRFIPKPDGLRPIYNMUYVGARFREREKRAERLTSVKA 682
Db 661 LFSVLTNERARRPGLGASVIGLDDIHRAMRTVLRVRAODPPPELTFVAVDTGADTI 720
QY 683 LFSVLTNERARRPGLGASVIGLDDIHRAMRTVLRVRAODPPPELTFVAVDTGADTI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVVQKAAGHVRAKAFSHYSTLTDLPYMQOFAHL 780
QY 743 PODRLTEVIASIIKPONTYCVRRYAVVQKAAGHVRAKAFSHYSTLTDLPYMQOFAHL 802
Db 781 QETSPFLDAVAVIESSSLNASSGLFDVFLRFPMCHAHAVRLRGSGSYVQCGGIIPOGSIITL 840

|||||
QY 803 QETSPRLDAVIEOSSSLNEASSGLFEDVLEFRFCHHAVIRKSKYQCGIPOGSIITSL 862
Db 841 LCSLCGDMENKIFAGIRRDGLLRVDVDFLVTPLHAKTFLRLNGVPEYGVNL 900
QY 863 LCSLCGDMENKIFAGIRRDGLLRVDVDFLVTPLHAKTFLRLNGVPEYGVNL 922
Db 901 RTVVPVEDEALGTAFAVQMPAGLFPACGLLDTRELEVOSSYARISASLTF 960
QY 923 RTVVPVEDEALGTAFAVQMPAGLFPACGLLDTRELEVOSSYARISASLTF 982
Db 961 NRGFKAGMRKRLFGVLRKCHSLFLDQVNSIQVCTNIRKILLQAVRFAVYQLP 1020
QY 983 NRGFKAGMRKRLFGVLRKCHSLFLDQVNSIQVCTNIRKILLQAVRFAVYQLP 1042
Db 1021 FHOQVKNPTFFLRVSDPASLCYSLKAKNAGMSLGAGAGAPLSEAVQVLCQAFL 1080
QY 1043 FHOQVKNPTFFLRVSDPASLCYSLKAKNAGMSLGAGAGAPLSEAVQVLCQAFL 1102
Db 1081 KLTRHRTVVPPLGSLRTAQTOLSRKLPGLTTLFALMAANPALPSDFKTIID 1132
QY 1103 KLTRHRTVVPPLGSLRTAQTOLSRKLPGLTTLFALMAANPALPSDFKTIID 1154
RESULT 15
ID US-08-912-951-2 STANDARD: PR1: 1132 AA.
XX xxxxxx
DE Sequence 2, Application US/08912951
XX
CC Sequence 2, Application US/08912951
CC GENERAL INFORMATION:
CC APPLICANT: Cecch, Thomas R.
CC APPLICANT: Lingner, Joachim
CC APPLICANT: Nakamura, Toru
CC APPLICANT: Chapman, Karen B.
CC APPLICANT: Morin, Gregg B.
CC APPLICANT: Harley, Calvin
CC APPLICANT: Andrews, William H.
CC TITLE OF INVENTION: HUMAN TELOMERASE CATALYTIC SUBUNIT: DIAGNOSTIC AND
CC TITLE OF INVENTION: THERAPEUTIC METHODS
CC NUMBER OF SEQUENCES: 335
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Townsend and Townsend and Crew LLP
CC STREET: Two Embarcadero Center, 8th Floor
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/912,951
CC FILING DATE: 14-AUG-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/854,050
CC FILING DATE: 09-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/851,843
CC FILING DATE: 06-MAY-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/846,017
CC FILING DATE: 25-APR-1997
CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/844,419
CC FILING DATE: 18-APR-1997
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/724,643
CC FILING DATE: 01-OCT-1996
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Apple, Randolph T.
CC REGISTRATION NUMBER: 36,429
CC REFERENCE/DOCKET NUMBER: 015389-002600US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 576-0200
CC TELEFAX: (415) 576-0300
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1132 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 1132 AA; 12695 MW; 6588209 CN;
SQ
Query Match 98.2%; Score 8465; DB 13; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 1132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 MPAPRCRAVSLNSHREVLPLATFVRRLGPGQMRVQGRDPAFALVACLVCPW 60
QY 23 MPAPRCRAVSLNSHREVLPLATFVRRLGPGQMRVQGRDPAFALVACLVCPW 82
Db 61 DAPPPAPSPFROVGLKELVAVLQRLCERGAKNVLAFGFALDGAAGCPREATTSVR 120
QY 83 DAPPPAPSPFROVGLKELVAVLQRLCERGAKNVLAFGFALDGAAGCPREATTSVR 142
Db 121 SYLPTVTDLALSGAWGILLRRVGDVYLHLARCAFLVLAPECAVQVCPPLYOLGA 180
QY 143 SYLPTVTDLALSGAWGILLRRVGDVYLHLARCAFLVLAPECAVQVCPPLYOLGA 202
Db 181 AFOARPPHAGSPRRRLGCEERAMNHSVREAGVPLGAPGARRGGSASRSJPLPKRPR 240
QY 203 AFOARPPHAGSPRRRLGCEERAMNHSVREAGVPLGAPGARRGGSASRSJPLPKRPR 262
Db 241 GAPEPERTPVQGSVAHAGRTGRGSDRGFCVSPAPRAPEATSLGALSGRRHSPSVG 300
QY 263 GAPEPERTPVQGSVAHAGRTGRGSDRGFCVSPAPRAPEATSLGALSGRRHSPSVG 322
Db 301 RQHHAGPSTSPRPWDTPCPVYAEETKHFLYSSGDKQLRPSFLSLRSLTGARL 360
QY 323 RQHHAGPSTSPRPWDTPCPVYAEETKHFLYSSGDKQLRPSFLSLRSLTGARL 382
Db 361 VETIFLGSPPMPGTPRRRLPQRYWQMRPLFELGNAHCOPYGLIKTHCPYRAVT 420
QY 383 VETIFLGSPPMPGTPRRRLPQRYWQMRPLFELGNAHCOPYGLIKTHCPYRAVT 442
Db 421 PAAGCAREKPGGSVAAPREEDTPRRVLQRLRQHSPPQVGVFACLRRLVPPGWS 480
QY 443 PAAGCAREKPGGSVAAPREEDTPRRVLQRLRQHSPPQVGVFACLRRLVPPGWS 502
Db 481 RNNERFLRNTKFTSLGKHAFLSLQELTWKMSVDCAMLRSPVGVCPAAEHRLREI 540
QY 503 RNNERFLRNTKFTSLGKHAFLSLQELTWKMSVDCAMLRSPVGVCPAAEHRLREI 562
Db 541 LAKFLHMLSVYVELLSSEFYVETTRQKNLFFRYKSVMSKLOSIGIRQHLKRYOLRE 600
QY 563 LAKFLHMLSVYVELLSSEFYVETTRQKNLFFRYKSVMSKLOSIGIRQHLKRYOLRE 622
Db 601 LSEAEVQRHREARPALTSRLRFIPKPSLRIVMMDVYVAGTRRREARRLTSRYVA 660
QY 623 LSEAEVQRHREARPALTSRLRFIPKPSLRIVMMDVYVAGTRRREARRLTSRYVA 682
Db 661 LPSVLYNERARPPGLLGASVGLDDIHRAMRTFVVRVQADPPPLTYFKVDVTGAYDTI 720

```

Oy 683 LFSVLNTERARRPGLIGASVLGLDDIHRAMTFVLNRADPPPELYEVKVDVTGAYDTI 742
Db 721 PODRLTEVIASIIKPONTYCVRRYAVNOKAAGHVKAFAKSHVSTLTDLOPYMRQFVAHL 780
Oy 743 PODRLTEVIASIIKPONTYCVRRYAVNOKAAGHVKAFAKSHVSTLTDLOPYMRQFVAHL 802
Db 781 QETSPLRDAVVIQOSSSLNEASSGLFDVFLRPMCHHAVIRGKSYVOCGIPQGSILSTL 840
Oy 803 QETSPLRDAVVIQOSSSLNEASSGLFDVFLRPMCHHAVIRGKSYVOCGIPQGSILSTL 862
Db 841 LCSLCYGDMEKFLFAGIRRDGILLRLVDDFLLVTPHLTHAKTFRLTVRGVPEYGCYVNL 900
Oy 863 LCSLCYGDMEKFLFAGIRRDGILLRLVDDFLLVTPHLTHAKTFRLTVRGVPEYGCYVNL 922
Db 901 RRTVYNFPEDEALGTAFAVQMPAHGLFPWCGLLDTRILEVQSDYSSYARTSIRASLTF 960
Oy 923 RRTVYNFPEDEALGTAFAVQMPAHGLFPWCGLLDTRILEVQSDYSSYARTSIRASLTF 982
Db 961 NRGFKAGRMRRKFLFVLRKCHSLFLDLQVNSIQVCTNIYKILLQAYRPHACVLOLP 1020
Oy 983 NRGFKAGRMRRKFLFVLRKCHSLFLDLQVNSIQVCTNIYKILLQAYRPHACVLOLP 1042
Db 1021 FHOQVWKNPTEFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLCQAFL 1080
Oy 1043 FHOQVWKNPTEFLRVISDTASLCYSILKAKNAGMSLGAKGAGPLPSEAVQWLCQAFL 1102
Db 1081 KTRRRVTVVPLLGSLRTAQOTLSRKLPGTTTLTALEAANPALPSDEKTIID 1132
Oy 1103 KTRRRVTVVPLLGSLRTAQOTLSRKLPGTTTLTALEAANPALPSDEKTIID 1154

```

Search completed: Fri Dec 18 18:34:40 1998
 Job time : 197 secs.

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 621,849
FILING DATE: 04-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 616,374
FILING DATE: 21-NOV-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 600,024
FILING DATE: 18-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 599,543
FILING DATE: 18-OCT-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 579,865
FILING DATE: 07-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 569,920
FILING DATE: 20-AUG-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 483,913
FILING DATE: 22-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 422,613
FILING DATE: 17-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 315,342
FILING DATE: 23-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 232,630
FILING DATE: 15-AUG-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 179,460
FILING DATE: 08-APR-1988
ATTORNEY/AGENT INFORMATION:
NAME: PITCHER, EDMUND R.
REGISTRATION NUMBER: 27,829
REFERENCE/DOCKET NUMBER: CRP-001CP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 402 AA; 44698 MW; 800499 CN;
Query Match 1.8%; Score 125; DB 1; Length 402;
Best Local Similarity 39.7%; Pred. No. 5,57e-01;
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;
Db 12 GLALGAGGGGGLRPP-CC-PQRLG-ARERDVQREILAVGLPGR-PRPRAPPAAS 67
QY 194 GPLLYQLGATQA-RPPHSGPRLRLGCEERAMNHSVREAGVPLGLPAGARRGGSASR 252
Db 68 RLP 70
QY 253 SLP 255
RESULT 2
ID US-08-643-763A-21 STANDARD; PRT; 402 AA.
AC xxxxxx
XX
XX
XX
XX
XX
DE Sequence 21, Application US/08643763A
XX
XX Sequence 21, Application US/08643763A
CC Patent No. 5733878
CC GENERAL INFORMATION:

APPLICANT: KUBERASAMPATH, THANANGVEL
APPLICANT: RUEGER, DAVID C.
APPLICANT: OPPERMAN, HERMAN
APPLICANT: COHEN, CHARLES M.
APPLICANT: PANG, ROY H.L.
TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE
TITLE OF INVENTION: REGENERATION.
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
STREET: 45 SOUTH STREET
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 01748
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/643,763A
FILING DATE: 06-MAY-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: FENTON Esq., GILLIAN M.
REGISTRATION NUMBER: 36,508
REFERENCE/DOCKET NUMBER: CRP-067CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 248-7560
TELEFAX: (617) 248-7100
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 402 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 402 AA; 44764 MW; 803627 CN;
Query Match 1.8%; Score 125; DB 1; Length 402;
Best Local Similarity 39.7%; Pred. No. 5,57e-01;
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;
Db 12 GLALGAGGGGGLRPP-CC-PQRLG-ARERDVQREILAVGLPGR-PRPRAPPAAS 67
QY 194 GPLLYQLGATQA-RPPHSGPRLRLGCEERAMNHSVREAGVPLGLPAGARRGGSASR 252
Db 68 RLP 70
QY 253 SLP 255
RESULT 3
ID US-08-451-953A-21 STANDARD; PRT; 402 AA.
AC xxxxxx
XX
XX
XX
XX
XX
DE Sequence 21, Application US/08451953A
XX
XX Sequence 21, Application US/08451953A
CC Patent No. 5741641
CC GENERAL INFORMATION:
CC APPLICANT: SMART, JOHN
CC APPLICANT: OPPERMAN, HERMAN
CC APPLICANT: OZKAYNAK, ENGİN
CC APPLICANT: KUBERASAMPATH, THANANGVEL
CC APPLICANT: RUEGER, DAVID C.
CC APPLICANT: PANG, ROY H.L.
CC APPLICANT: COHEN, CHARLES M.
CC TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD
CC NUMBER OF SEQUENCES: 33

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/451,953A
CC FILING DATE: 26-MAY-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: PITCHER ESQ., EDMUND R.
CC REGISTRATION NUMBER: 27,829
CC REFERENCE/DOCKET NUMBER: CRP-058CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508) 435-9001
CC TELEFAX: (508) 435-6951
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 402 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 402 AA; 44764 MW; 803627 CN;
SQ

Query Match 1.8%; Score 125; DB 1; Length 402;
Best Local Similarity 39.7%; Pred. No. 5,57e-01;
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;

Db 12 GIALCALGGGGPGLRPP-GC-PQRLG-ARERDVQREILLAVGLPGR-PPRPAPPAAS 67
QY 194 GPPLYLQGLAAQA-RPPHASGPRRLGCERAMNHSVREAGVPLGLPAPGARRRGGSASR 252
DB 68 RLP 70
QY 253 SLP 255

RESULT 4
ID US-08-462-623-21 STANDARD; PRT; 402 AA.
AC xxxxxx
XX
DT
DE
XX Sequence 21, Application US/08462623
XX Sequence 21, Application US/08462623
CC Patent No. 5739107
CC GENERAL INFORMATION:
CC APPLICANT: COHEN, CHARLES M.
CC APPLICANT: CHARETTE, MARC F.
CC APPLICANT: KUBERASAMPATH, THANGAVEL
CC APPLICANT: RUEGER, DAVID C.
CC APPLICANT: OPPERMAN, HERMANN
CC APPLICANT: PANG, ROY H.L.
CC APPLICANT: OKRAYNAK, ENGIN
CC APPLICANT: SMART, JOHN E.
CC TITLE OF INVENTION: MORPHOGEN TREATMENT OF GASTROINTESTINAL
CC TITLE OF INVENTION: ULCERS.
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA

CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/462,623
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/445,882
CC FILING DATE: 22-MAY-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FENTON ESQ., GILLIAN M.
CC REGISTRATION NUMBER: 36,508
CC REFERENCE/DOCKET NUMBER: CRP-074CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (508) 435-9001
CC TELEFAX: (508) 435-6951
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 402 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 402 AA; 44764 MW; 803627 CN;
SQ

Query Match 1.8%; Score 125; DB 1; Length 402;
Best Local Similarity 39.7%; Pred. No. 5,57e-01;
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;

Db 12 GIALCALGGGGPGLRPP-GC-PQRLG-ARERDVQREILLAVGLPGR-PPRPAPPAAS 67
QY 194 GPPLYLQGLAAQA-RPPHASGPRRLGCERAMNHSVREAGVPLGLPAPGARRRGGSASR 252
DB 68 RLP 70
QY 253 SLP 255

RESULT 5
ID PCT-US93-08742-21 STANDARD; PRT; 402 AA.
AC xxxxxx
XX
DT
DE
XX Sequence 21, Application PC/TUS9308742
XX Sequence 21, Application PC/TUS9308742
CC GENERAL INFORMATION:
CC APPLICANT:
CC TITLE OF INVENTION: MORPHOGEN-INDUCED PERIODONTAL TISSUE REGENERATION
CC NUMBER OF SEQUENCES: 33
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.
CC STREET: 45 SOUTH STREET
CC CITY: HOPKINTON
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 01748
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US93/08742
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:

```

CC      FILING DATE:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: KELLEY ESQ, ROBIN D.
CC      REGISTRATION NUMBER: 34,637
CC      REFERENCE/DOCKET NUMBER: CRP-067
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 617/248-7477
CC      TELEFAX: 617/248-7100
CC      INFORMATION FOR SEQ ID NO: 21:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 402 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 402 AA; 44698 MW; 800499 CN;
CC
CC
CC      Query Match          1.8%; Score 125; DB 2; Length 402;
CC      Best Local Similarity 39.7%; Pred. NO. 5,57e-01;
CC      Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5
CC
CC      Db      12 GLALCALGGGPGCLRPP-GC-PQRRLG-AERRDVQREITAVLGLPGR-PRPAPPAAS 67
CC      QY      194 GPLPVLQGAATAA-RPPHSGPFRRLGCEERAMNHSYREAGVPLGLPAPGARRGGSASR 252
CC      Db      68 RLP 70
CC      QY      253 SLP 255
CC
CC      RESULT      6          STANDARD:      PRT;      402 AA.
CC      ID      US-08-206-864-4
CC      XX      xxxxxx
CC      DT
CC      XX
CC      DE      Sequence 4, Application US/08206864
CC      XX
CC      Sequence 4, Application US/08206864
CC      Patent No. 5610021
CC      GENERAL INFORMATION:
CC      APPLICANT: RUEGER, DAVID C
CC      APPLICANT: JONES, WILLIAM K
CC      APPLICANT: TUCKER, RONALD F
CC      APPLICANT: OPPERMAN, HERMANN
CC      APPLICANT: OZKAYNAK, ENGIN
CC      APPLICANT: KUBERASAMPATH, TANGAVEL
CC      TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR
CC      TITLE OF INVENTION: RECOMBINANT OSTEOGENIC PROTEIN PRODUCTION
CC      NUMBER OF SEQUENCES: 6
CC      CORRESPONDENCE ADDRESSES:
CC      ADDRESSEE: PATENT ADMINISTRATOR/CREATIVE BIOMOLECULES,
CC      ADDRESSEE: INC.
CC      STREET: 45 SOUTH STREET
CC      CITY: HOPKINTON
CC      STATE: MA
CC      COUNTRY: USA
CC      ZIP: 01748
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/206,864
CC      FILING DATE:
CC      CLASSIFICATION: 530
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 08/027,070
CC      FILING DATE: 04-MAR-1993
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/841,646
CC      FILING DATE: 21-FEB-1992
CC

```

```
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: PITCHER, EDMOND R.  
CC REGISTRATION NUMBER: 27, 829  
CC REFERENCE/DOCKET NUMBER: CRP-096  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617//7000  
CC TELEFAX: 617/248-7100  
CC INFORMATION FOR SEQ ID NO: 4:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 402 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 402 AA; 44698 MW; 800499 CN;  
SQ  
  
Query Match 1.8%; Score 125; DB 1; Length 402;  
Best Local Similarity 39.7%; Pred. No. 5.57e-01;  
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;  
  
Db 12 GLALCALGGGPGRLRPP-GC-PQRRLG-AERRDVQRELIATVLGLGR-PRRPAPPAAS 67  
QY 194 GPLPLQGAAQTQA-RPPHAGSPRRRLGCERAMNHSVEAGVPGLGPAGARRRGGSASR 252  
Dy 68 RLP 70  
QY 253 SLP 255  
  
RESULT 7 STANDARD; PRT; 402 AA.  
ID US-08-155-343A-21  
XX xxxxxx  
DT  
XX  
XX  
XX  
XX  
Sequence 21, Application US/08155343A  
DE  
CC Sequence 21, Application US/08155343A  
CC Patent No. 5656593  
CC GENERAL INFORMATION:  
CC APPLICANT: KUBERASAMPATH, THANGAVEL  
CC APPLICANT: RUEGER, DAVID C.  
CC APPLICANT: OPPERMAN, HERMAN  
CC APPLICANT: COHEN, CHARLES W.  
CC APPLICANT: PANG, ROY H. L.  
CC TITLE OF INVENTION: MORPHOGENIC-INDUCED PERIODONTAL TISSUE  
CC TITLE OF INVENTION: REGENERATION.  
CC NUMBER OF SEQUENCES: 33  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES  
CC STREET: 45 SOUTH STREET  
CC CITY: HOPKINTON  
CC STATE: MA  
CC ZIP: 01748  
CC COUNTRY: USA  
CC  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patentin Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/155,343A  
CC FILING DATE: 15-NOV-1993  
CC CLASSIFICATION: 435  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: FEWTON ESQ., GILLIAN M.  
CC REGISTRATION NUMBER: 36,508  
CC REFERENCE/DOCKET NUMBER: CRP-067FW  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617) 248-7100  
CC TELEFAX: (617) 248-7100  
CC INFORMATION FOR SEQ ID NO: 21:  
CC SEQUENCE CHARACTERISTICS:  
CC
```

```
CC      LENGTH: 402 amino acids  
CC      TYPE: amino acid  
CC      TOPOLOGY: linear  
CC      MOLECULE TYPE: protein  
SC      SEQUENCE 402 AA; 44764 MW; 803627 CN;  
  
Query Match                               1.8%; Score 125; DB 1; Length 402;  
Best Local Similarity 39.7%; Pred. No. 5,57e-01;  
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5.  
  
Db     12 GLALCALGGGPGRLRPP-GC-PQRRLG-AERRDVGREILAVGLDGR-PRRAPPAAS 67  
OY     194 GPPLTQLGAATAQA-RPPHAGSPRRRLGCCERAMNHVSVEAGVPGLGPAPGARRRRGSSASR 252  
DB     68 RLP 70  
OY     253 SLP 255  
  
RESULT   8 STANDARD; PRT; 402 AA.  
ID PC1-US93-07231-21  
XX AC xxxxxx  
DX DT  
XX AC  
XX AC  
XX AC  
DE Sequence 21, Application PC/TUS9307231  
CX Sequence 21, Application PC/TUS9307231  
CC GENERAL INFORMATION:  
CC APPLICANT:  
CC TITLE OF INVENTION: MORPHOGEN-INDUCED NERVE REGENERATION AND  
CC TITLE OF INVENTION: REPAIR  
CC NUMBER OF SEQUENCES: 33  
CC CORRESPONDENCE ADDRESS:  
CC ADDRESSEE: CREATIVE BIOMOLECULES, INC.  
CC STREET: 35 SOUTH STREET  
CC CITY: HOPKINTON  
CC STATE: MASSACHUSETTS  
CC COUNTRY: USA  
CC ZIP: 01748  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: Patent In Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT/US93/07231  
CC FILING DATE: 19930729  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: KELLEY, ROBIN D.  
CC REGISTRATION NUMBER: 34,637  
CC REFERENCE/DOCKET NUMBER: CRP-070  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: 617/248-7000  
CC TELEX: 617/248-7100  
CC INFORMATION FOR SEQ ID NO: 21:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 402 amino acids  
CC type: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: protein  
CC SEQUENCE 402 AA; 44698 MW; 800499 CN;
```

[illegible]

CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 402 amino acids
CC	TYPE: amino acid
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: protein
SQ	SEQUENCE 402 AA; 44698 MW; 800499 CN;
Db	Query Match 1.8%; SCORE 125; DB 1; Length 402;
QY	Best Local Similarity 39.7%; Pred.No.5.57e-01;
	Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;
Db	12 GLALCALGGSGPRPP-GC-PORRLG-ARERDVOREILAVLGLPGR-PRPRAPPAS 67
QY	194 GPLYLQGAIAQA-RPPPHASGPRRLRIGCEKAMNHSYREAGVPLGLPAPGARRRGGSASR 252
Db	68 RLP 70
QY	253 SLP 255
RESULT 12	STANDARD; PRT; 402 AA.
ID	US-08-643-563A-21
XX	xxxxxx
XX	
DT	
DE	Sequence 21, Application US/08643563A
XX	
CC	Sequence 21, Application US/08643563A
CC	Patent No. 5707810
CC	GENERAL INFORMATION:
CC	APPLICANT: SMART, JOHN
CC	APPLICANT: OPPERMAN, HERMAN
CC	APPLICANT: OZKAYNAK, ENGIN
CC	APPLICANT: KUBERASAMPATH, THANGAVEL
CC	APPLICANT: RUEGER, DAVID C.
CC	APPLICANT: PANG, ROY H.L.
CC	APPLICANT: COHEN, CHARLES M.
CC	TITLE OF INVENTION: MORPHOGENIC PROTEIN SCREENING METHOD
CC	NUMBER OF SEQUENCES: 33
CC	CORRESPONDENCE ADDRESS:
CC	ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES
CC	STREET: 45 SOUTH STREET
CC	CITY: HOPKINTON
CC	STATE: MA
CC	COUNTRY: USA
CC	ZIP: 01748
CC	COMPUTER READABLE FORM:
CC	MEDIUM TYPE: Floppy disk
CC	COMPUTER: IBM PC compatible
CC	OPERATING SYSTEM: PC-DOS/MS-DOS
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30
CC	CURRENT APPLICATION DATA:
CC	APPLICATION NUMBER: US/08/643,563A
CC	FILING DATE: 06-MAY-1996
CC	CLASSIFICATION: 435
CC	ATTORNEY/AGENT INFORMATION:
CC	NAME: TWOMEY ESQ., MICHAEL J.
CC	REGISTRATION NUMBER: 38,349
CC	REFERENCE/DOCKET NUMBER: CRP-058CN2
CC	TELECOMMUNICATION INFORMATION:
CC	TELEPHONE: (508) 435-9001
CC	TELEFAX: (508) 435-6951
CC	INFORMATION FOR SEQ ID NO: 21:
CC	SEQUENCE CHARACTERISTICS:
CC	LENGTH: 402 amino acids
CC	TYPE: amino acid
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: protein
SQ	SEQUENCE 402 AA; 44764 MW; 803627 CN;
Query Match	1.8%; SCORE 125; DB 1; Length 402;

BEST LOCAL SIMILARITY	39.7%	PRED. NO.	5.57e-01;
MATCHES	25;	CONSERVATIVE	12; MISMATCHES 21; INDELS 5; GAPS 5;
Db	12	GLATCALGGGPGRLPPP-GC-PQRRLG-ARERDVGRELLAVLGLDGR-PPRAPPAAAS	67
Oy	194	GPLYVLQAATAA-QRPPhASGPFRRLGCERAMNHSYREAGVPLGLPAPGARRRGSASR	252
Dd	68	RLP 70	
Oy	253	SLP 255	
RESULT	13	STANDARD:	PRT; 402 AA.
ID	US-08-406-672-21		
AC	xxxxxx		
DT			
DE	Sequence 21, Application US/08406672		
XX			
CC	Sequence 21, Application US/08406672		
CC	Patent No. 56744844		
CC	GENERAL INFORMATION:		
CC	APPLICANT: KUBERASAMPATH, THANCAVEL		
CC	APPLICANT: COHEN, CHARLES M.		
CC	APPLICANT: OPPERMAN, HERMANN		
CC	APPLICANT: OKRAYNAK, ENGIN		
CC	APPLICANT: RUEGER, DAVID C.		
CC	APPLICANT: PANG, ROY H.L.		
CC	TITLE OF INVENTION: TREATMENT TO PREVENT LOSS OF AND/OR		
CC	TITLE OF INVENTION: INCREASE BONE MASS IN METABOLIC BONE DISEASES		
CC	NUMBER OF SEQUENCES: 33		
CC	CORRESPONDENCE ADDRESSES:		
CC	ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES		
CC	STREET: 45 SOUTH STREET		
CC	CITY: HOPEKINTON		
CC	STATE: MA		
CC	COUNTRY: USA		
CC	ZIP: 01748		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.30		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/406,672		
CC	FILING DATE: 20-MAR-1995		
CC	CLASSIFICATION: 435		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 752,857		
CC	FILING DATE: 30-AUG-1991		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: US 667,274		
CC	FILING DATE: 11-MAR-1991		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: FENTON ESQ., GILLIAN M.		
CC	REGISTRATION NUMBER: 36,508		
CC	REFERENCE/DOCKET NUMBER: CRP-0600CN		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: (617) 248-7560		
CC	TELEFAX: (617) 248-7100		
CC	INFORMATION FOR SEQ ID NO: 21:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 402 amino acids		
CC	TYPE: amino acid		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: protein		
SO	SEQUENCE 402 AA; 44764 MW; 803627 CN;		
Query Match	1.8%; Score 125; DB 1; Length 402;		
Best Local Similarity	39.7%; Pred. No. 5.57e-01;		
Matches	25; Conservative 12; Mismatches 21; Indels 5; Gaps 5;		

MUSEUM
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPERCH_MP protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:10:34 1998; Maspar time 33.93 Seconds
954.969 Million cell updates/sec
Tabular output not generated.

Title: >US-08-951-733-14
Description: (1-949) from US08951733.pep (1 of 2)
Perfect Score: 7113

Sequence: 1 HASGQRCVLTFTWEALAPAT.....PVEDEALGTAFTVOMPAHGL 949

Scoring table:
PAM 150
Gap 11

Searched: 107076 seqs, 34141958 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

pir56
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 53.200; Variance 116.643; scale 0.456

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	314	4.4	884	2	S53396	hypothetical protein	1.95e-30
2	162	2.3	660	1	Q08B3	BHLFI protein - human	6.26e-07
3	153	2.2	504	2	JC1306	virion protein homolo	1.03e-05
4	141	2.0	276	2	B38965	hypothetical protein	3.83e-04
5	142	2.0	310	1	PIHUSD	saliary prolina-rich	2.85e-04
6	136	1.9	240	2	B24264	proline-rich protein	1.65e-03
7	136	1.9	316	2	S16681	homeotic protein - hu	1.65e-03
8	132	1.9	317	2	A28996	proline-rich protein	1.19e-03
9	137	1.9	628	2	S01955	hypothetical protein	1.23e-03
10	132	1.9	924	2	S27923	gene Lf3 protein - hu	5.19e-03
11	129	1.8	227	2	C29149	proline-rich protein	3.67e-02
12	129	1.8	264	2	D34768	ORF4 protein - Orf v1	1.21e-02
13	131	1.8	300	2	S19560	proline-rich protein	6.89e-03
14	126	1.8	302	2	S11790	nodulation protein no	2.79e-02
15	129	1.8	309	2	S10889	proline-rich protein	1.21e-02
16	125	1.8	381	2	S16506	hypothetical protein	3.67e-02
17	125	1.8	402	2	A45056	prepro osteogenic pro	3.67e-02
18	127	1.8	566	2	S22933	testis-specific prote	2.11e-02
19	131	1.8	1106	2	JQ0405	hypothetical 119.5K p	6.89e-03
20	123	1.7	240	2	A24264	proline-rich protein	6.35e-02
21	123	1.7	256	2	A60533	tumor-associated anti	6.35e-02
22	118	1.7	332	1	S00054	nodulation protein no	2.43e-01
23	119	1.7	347	2	S10571	epithelial tumor anti	1.87e-01

24	121	1.7	358	1	MMB38	infected cell protein	1.09e-01
25	123	1.7	403	2	S52796	prp2 protein - human	6.35e-02
26	118	1.7	431	2	S09824	hypothetical protein	2.43e-01
27	123	1.7	439	2	S51939	chitinase (EC 3.2.1.1	6.35e-02
28	124	1.7	464	2	S22697	extensin - Volvox car	4.83e-02
29	119	1.7	515	2	S10572	epithelial tumor anti	1.87e-01
30	114	1.6	202	2	B36795	hypothetical protein	6.93e-01
31	117	1.6	204	2	A39066	proline-rich protein	3.17e-01
32	115	1.6	260	2	S22373	proline-rich protein	5.34e-01
33	116	1.6	270	2	S34361	mlaf protein - Salmon	4.12e-01
34	116	1.6	295	2	B48013	proline-rich proteogl	4.12e-01
35	114	1.6	301	2	E29149	proline-rich protein	6.93e-01
36	114	1.6	321	1	A28663	nodulation protein no	6.93e-01
37	115	1.6	338	2	JC5407	HXA2 protein - human	5.34e-01
38	114	1.6	373	2	S54545	hypothetical protein	6.93e-01
39	115	1.6	392	1	PIHUB6	saliary prolina-rich	5.34e-01
40	114	1.6	408	2	A56186	cyclin E - African cl	6.93e-01
41	113	1.6	668	2	B35363	synapsin Ib - human	8.96e-01
42	113	1.6	704	2	A30411	synapsin Ia - rat	8.96e-01
43	114	1.6	705	2	A35363	synapsin I splice for	6.93e-01
44	113	1.6	1335	2	A35887	mucin precursor, panc	8.96e-01
45	115	1.6	3149	1	Q08B8	BPLFI protein - human	5.34e-01

ALIGNMENTS

RESULT ENTRY TITLE	1	S53396	#type complete
		hypothetical protein YLR318w - yeast (Saccharomyces cerevisiae)	

ALTERNATE_NAMES ORGANISM DATE	hypothetical protein L8543.12	#formal_name Saccharomyces cerevisiae	#sequence_revisions 01-Sep-1995	#text_change
		05-May-1995		
		06-Feb-1998		

ACCESSIONS S53396

REFERENCE S53390

#authors

#submitted to the EMBL Data Library, February 1995

#description The sequence of S. cerevisiae cosmid 8543.

#accession S53396

#molecule_type DNA

#residues 1-884

#cross-references EMBL:U00618; NID:g2258165; PID:g662136; MIPS:YLR318w

#experimental_source strain S288C (AB972)

GENETICS

#gene

#cross-references SGD:S0004310; MIPS:YLR318w

#map_position 12R

SUMMARY

#length 884

#molecular-weight 102662

#checksum 7604

Query Match

Best Local Similarity 24.3%; Pred. No. 1.95e-30;

Matches 118; Conservative 122; Mismatches 202; Indels 44; Gaps 37;

Db

271 LSHLSKSPKPR-VLFTIVIIQKLPDQMSKKNKGIINNLISLPUNGIPFSPS 329

QY

470 LVQLNQHSSPMQVGVFACRLRPGLMGRNHRERFLNTKIFSLGHAKLSDE 529

Db

330 LKKLKLKDFMLFIS-DLWFTKHNEFNOLAIC-FISMLRQILPKIIOFFCTEIS 387

QY

530 LTKKSVKRCANLRSPPGVCYPAEHRRLRETLAKFTLWMLSVYVELRFFVTEYT 589

Db

388 STVTIYVF-RHDTWNLITPFTVEYFKY-LVE-NNV-CRNHSYTLNFSNKKMIIRK 443

QY

590 FQKNRLFFYKRSWMSKLSIGIRHKLRYQLBELSBAEYORRE-ARPALTSRLRIFPK 648

Db

444 KSNERRIITAICRGDEDEFTIYKENR-NA-IQPTQILILYLNKRRTSF-TTIYST 500

QY

649 -PDGLRPVNMXY-VVGARTFRERKAERLTSRVALFSLVYERARRGGLGASVTLGD 706

Db

501 QIADRI-KEFKQLLKKFNUNVPELYFMKFDVSKSDSPRRECKR-IKDKALKNNGFF 558

QY

707 DLRHAKRTFVLRV-RA-QDPEELTFYVVDYTGADITIPQ-DRLTEVYASITIKFQNTIC 762

Db	559	VR50Y-FENNTM-G-VLKLFW-VWASVPRKY-ELYIDNVR-TVHLSNDVIN---VV-	608
Oy	763	VR-KRAVQKAAGHGVKRAKFSHVSITLDDLPYMRQFYAHQLQETSPLRDAVYIOSSLN	821
Db	609	E---M-EIF-K-T---ALWEDKRCYIREDELFGSSLSAPIVDYVDLLEFYSEFKAS	658
Oy	822	EASSGLFDFVRFMKHNAVIRIRGSKSYVQCQIGIPOGSIITSLCISLGYDM-E-NKLF-AG	878
Db	659	PSQDPLIKLADDFLIISTDOQVYINIKKLAMGEGFQKNAAN-RDKILAVSSSDDDTV	717
Oy	879	IRRDGLLRIVLDFELVPLHLTHAKFTLRTLVRGVPEYGVVNAKRTVYVNEPVEDALGG	938
Db	718	IQFCAM 723	
Oy	939	TAFVQM 944	
RESULT	2		
ENTRY		Q0B53	#type complete
TITLE		BHLF1 protein - human herpesvirus 4 (strain B95-8)	
ORGANISM		#formal name human herpesvirus 4, Epstein-Barr virus	
DATE		25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 23-Aug-1997	
ACCESSIONS		A03742	
REFERENCE		A93065	
#authors		Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.	
#journal		Mol. Biol. Med. (1983) 1:21-45	
#title		Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus.	
#cross-references		MUID:85035713	
#accession		A03742	
#molecule_type		DNA	
#residues		1-660	#label BAN
REFERENCE		A03794	
#authors		Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; Hatfull, G.; Hudson, G.S.; Satchwell, S.C.; Seguin, C.; Tuffnell, P.S.; Barrell, B.G.	
#journal		Nature (1984) 310:207-211	
#title		DNA sequence and expression of the B95-8 Epstein-Barr virus genome.	
#cross-references		MUID:84270667	
#contents		annotation: protein coding region	
COMMENT		The sequence contains four perfect repeats (residues 145-273, 274-398, 399-523, and 524-648).	
CLASSIFICATION		#superfamily human herpesvirus 4 BHLF1 protein	
SUMMARY		#length 660 #molecular_weight 66244 #checksum 89300	
Query Match		2.3%; Score 162; DB 1; Length 660;	
Best Local Similarity		30.8%; Pred. No. 6,266-07;	
Matches		48; Conservative 34; Mismatches 60; Indels 14; Gaps 12;	
Db	251	GPPPTSGAANO-RTHRRPDCPSRANRPGCRPMRR--R-SGAKRQHRPPGAQRSGP	306
Oy	194	GPPYQDAATQARPPRIASG-PR-R-LGGERAMNSHVRAGVPLGLIPAGA-KRRGGS	249
Db	307	TGGRPAAGAGCTPAAPPGGGAATPSPERGSGPADP--AAARLPRKEQERLP	364
Oy	250	ASRSRLPLKRRRRRGAAPRER-TVYGCGSMHPRGTRGSPDRGTCVVSAPARAEARSL	308
Db	365	QDLAAQRC-PAGPPPTSGA-AAQRTNR-RPPCP	397
Oy	309	GALSGTRHSHPSVGRQNHAGCPSTSRPPRPWDTPCP	344
RESULT	3		
ENTRY		JC1306	#type complete
TITLE		virion protein homolog - bovine herpesvirus 1	
ALTERNATE_NAMES		alpha TIF; BHV-1 protein homolog; ICP25; Vmw65; VP16	
ORGANISM		#formal name bovine herpesvirus 1	
DATE		05-Mar-1993 #sequence_revision 05-Mar-1993 #text_change 09-Sep-1997	
ACCESSIONS		JC1306; S24229	

REFERENCE	JC1306
#authors	Carpenter, D.E.; Mistr, V.
#journal	Gene (1992) 119:25-263
#title	Sequences of the bovine herpesvirus 1 homologue of herpes simplex virus type-1 alpha-trans-inducing factor (UL48).
#accession	JC1306
##molecule_type	DNA
##residues	1-504 ##label CAR
##cross-references	EMBL:Z11610; NID:g1065725; PID:c264419; PID:g1065726
COMMENT	This protein interacts with cellular transcription factors to transactivate immediate early viral genes.
GENETICS	
##map_position	0.07-0.086
KEYWORDS	DNA binding; transcription regulation
SUMMARY	#length 504 #molecular-weight 54028 #checksum 8743
Query Match	2.2%; Score 153; DB 2; Length 504;
Best Local Similarity	31.3%; Pred. No. 1.03e-05;
Matches	46; Conservative 33; Mismatches 56; Indels 12; Gaps 11;
Db	351 APARAGGMRSSSTTRGAARSTTGRQRCCPRRRAC-CHAP-RRRLR-ARGE 406
Qy	208 PPPHASPRRRLCCERAMNHSVEAEVPLGLPAPGARRGSGASRSLPLPRRGAPE 267
Db	407 PRHNS-SSGAFSFG-GRPPGVNCLGWACAKARSGPARGGPPVSGSLGR-ARGSGCP 463
Qy	268 PERFPVQSGSWAHPGKTRGSPDR-GE-CVY-S-PARPAEATSLGALSGTRHSPVGR 323
Db	464 GPACGSPRARGGRRRAPAMP-FGGT 489
Qy	324 QHHAGPSTSRPPRPMDTPCPYVAT 350
RESULT	4
ENTRY	B38965
TITLE	hypothetical protein B (insertion sequence IS1222) -
ORGANISM	Enterobacter agglomerans
REFERENCE	#formal_name Enterobacter agglomerans
DATE	03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 09-Sep-1997
ACCESSIONS	B38965; S42922
REFERENCE	A38965
#authors	Steibl, H.D.; Lewecke, F.M.
#journal	Gene (1995) 156:37-42
#title	IS1222: analysis and distribution of a new insertion sequence in Enterobacter agglomerans 339.
#accession	B38965
##status	preliminary
##molecule_type	DNA
##residues	1-276 ##label STE
##cross-references	GB:X78052; NID:g459246; PID:g459248
SUMMARY	#length 276 #molecular-weight 31718 #checksum 3927
Query Match	2.0%; Score 141; DB 2; Length 276;
Best Local Similarity	33.7%; Pred. No. 3.83e-04;
Matches	30; Conservative 23; Mismatches 30; Indels 6; Gaps 6;
Db	42 ITTELALERRR-FEYRR-IMQLLRRELIHNHRVRYLYHLSGLGYKRRR-RKGIATERL 98
Qy	585 VTETTFCKNKLFFRYKRSVMSKLOSTIGIRHKLRY-QLEKLSAEVROHREARPALLSRL 643
Db	99 PLL-RPAAPLWSDDFVMDALATGRRIK 126
Qy	644 RFIPKPDGLRPIVNDYVGA--RTFRRER 671
RESULT	5
ENTRY	PIHUSD
TITLE	salivary proline-rich glycoprotein precursor PRB4 (large allele) - human
CONTAINS	basic proline-rich protein IB-5; proline-rich peptide P-D
DATE	19-Feb-1984 #sequence_revision 12-Apr-1996 #text_change

05-Sep-1997
S01175; S031175; S10890; D25372; E38355; A03295; A61294;
S62891

REFERENCE
#authors Lyons, K.M.; Stein, J.H.; Smithies, O.
#journal Genetics (1988) 120:267-278
#title Length polymorphisms in human proline-rich protein genes generated by intragenic unequal crossing over.
#cross-references MIM:89121440
#accession S03176
##status translation not shown
##molecule_type DNA
##residues 35-310 ##label LY1
##cross-references EMBL:X07715
##note large allele
#accession S03175
##status translation not shown
##molecule_type DNA
##residues 35-36, 'E', 38-112, 155-310 ##label LY2
##cross-references EMBL:X07704
##note medium allele
#accession S10890
##status preliminary; translation not shown
##molecule_type DNA
##residues 1-39, 60-112, 'I', 114-115, 'P', 117-121, 185-271, 'A', 273-310
##label LY3
#cross-references EMBL:X07882; MID:935647; PID:9296670
#accession A92492
#authors Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
#journal J. Biol. Chem. (1985) 260:11123-11130
#title Differential RNA splicing and post-translational cleavages in the human salivary proline-rich protein gene system.
#cross-references MIM:85289325
#accession D25372
##molecule_type mRNA
##residues 1-36, 'E', 38-112, 'I', 114-115, 'P', 117-121, 185-271, 'A', 273-310 ##label MAE

REFERENCE
#authors Kaufman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
#journal Biochemistry (1991) 30:3351-3356
#title Basic proline-rich proteins from human parotid saliva: relationships of the covalent structures of ten proteins from a single individual.
#cross-references MIM:9119084
#accession E38355
##molecule_type protein
##residues 241-254, 'KN', 257-310 ##label KAU

REFERENCE
#authors Salton, E.; Isemura, S.; Sanada, K.
#journal J. Biochem. (1983) 93:495-502
#title Complete amino acid sequence of a basic proline-rich peptide, P-D, from human parotid saliva.
#cross-references MIM:83186122
#accession A03295
##molecule_type protein
##residues 241-310 ##label SAI

REFERENCE
#authors Shimomura, H.; Kanai, Y.; Sanada, K.
#journal J. Biochem. (1983) 93:857-863
#title Amino acid sequences of glycopeptides obtained from basic proline-rich glycoprotein of human parotid saliva.
#accession A61294
##molecule_type protein
##residues 54-57, 'E', 59-73, 'R', 82-101 ##label SHI

REFERENCE
#authors Chailion, A.J.; Baxter, N.J.; Lilley, T.H.; Haslam, E.; McDonald, C.J.; Williamson, M.P.
#journal FEBS Lett. (1996) 382:289-292
#title Tannin interactions with a full-length human salivary proline-rich protein display a stronger affinity than with single proline-rich repeats.
#accession S62891
##molecule_type protein

##residues	241-252 ##label CHA
##note	amino end of peptide designated basic proline-rich protein IB-5
##note	it is unclear from the peptide sequence whether this is a product of the PRB2 (PIR:PIHUPF) or PRB4 (this entry) gene
GENETICS	
#gene	GDB:PRB4
##cross-references	GDB:119514; OMIM:180990
#map_position	12p13.2-12p13.2
#intons	22/1; 34/1
#note	the list of introns may be incomplete
CLASSIFICATION	#superfamily proline-rich protein
KEYWORDS	glycoprotein; saliva; tandem repeat
FEATURE	
1-16	#domain signal sequence #status predicted #label SIG\
241-310	#product proline-rich peptide P-D #status experimental
66,87,171	#label MAT\
108,150,192,213,	#binding site carbohydrate (Asn) (covalent) #status experimental\
234	#binding_site carbohydrate (Asn) (covalent) #status predicted
SUMMARY	#length 310 #molecular-weight 31351 #checksum 3960
Query Match	2.0%; Score 142; DB 1; Length 310;
Best Local Similarity	28.2%; Pred. No.2.85e-04;
Matches	40; Conservative 41; Mismatches 52; Indels 9; Gaps 9;
Db	168 QGNGSQG-P-PPHKGPERPP-OGG-NOSHRPPPP-GKPER-PPQGGNGSQG-P-PP 220
Qy	199 QGAATQARPPPHASGRRRLGGERAMNHSVREAGVPLGAPAGARRGGASASRLPLK 258
Db	221 HPGKRGPPROGGSRSKASRPPKQSGPPODGNKPPGPPGKQSGPPPPGNGNQPP 280
Qy	259 RPRRGAPRRPRPYGQGSWMHGRKSPSDR-GFCVSPAPAE-EATSLGSLGSTRH 316
Db	281 APPAGKPPGPPPPGPPGPP 302
Qy	317 SHPSVGRQHNAGRPSTRRPPR 338
RESULT	6
ENTRY	B24264 #type fragment
TITLE	proline-rich protein MP3 - mouse (fragment)
ORGANISM	#normal_name Mus musculus #common_name house mouse
DATE	09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 03-May-1996
ACCESSIONS	B24264
REFERENCE	A92508
#authors	Ann, D.K.; Carlson, D.M.
#journal	J. Biol. Chem. (1985) 260:15863-15872
#title	The structure and organization of a proline-rich protein gene of a mouse model gene family.
#cross-references	MUID:86059475
#accession	B24264
#molecule	DNA
CLASSIFICATION	#superfamily proline-rich protein
SUMMARY	#length 240 #checksum 5152
Query Match	1.9%; Score 136; DB 2; Length 240;
Best Local Similarity	27.2%; Pred. No.1.65e-03;
Matches	41; Conservative 38; Mismatches 65; Indels 7; Gaps 6;
Db	24 VNGSQGPPPPGPPQPPRPPGPPGPPQPPRPPGPPGPPGPPQPPRPPGPPGPPQPPR 83
Qy	200 LGATQARPPPHASGRRRLGGERAMNHSVREAGVPLGAPAGARRGGASASRLPLK 259
Db	84 PGGPPPPGPPQPPR-PGGPPPPGPPQPPRPPGPPGPPQPPRPPGPPGPPGPPGPP 142
Qy	260 PRGGAPE-PEKPPVQGGSMHNG-RTRGSPDGFVCSAPARAELTSLGGA-LSGTRH 316


```

#authors
Parker, B.D.; Bankier, A.; Satchwell, S.; Barrell, B.;
Barrell, P.J.
#submission
submitted to the EMBL Data Library, August 1990
#description
Sequence and transcription of Raj1 Epstein-Barr virus DNA
spanning the B95-8 deletion region.
#accession
S27923
##status
Preliminary
##molecule_type
DNA
##residues
1-924 ##label
PAR
##cross-references
EMBL:M35547; NID:g330420; PID:g330421
SUMMARY
length 924 #molecular_weight 94304 #checksum 8997

Query Match 1.9%; Score 132; DB 2; Length 924;
Best Local Similarity 28.2%; Pred. No. 5,19e-03;
Matches 37; Conservative 33; Mismatches 52; Indels 9; Gaps 9;

Db 38 AAPRAGPEPTRLQPTPRRSGAADPADPVGHPA-PRAPGEPTRLQ-PATPRRSGA 95
| : : : | | : : : : : : : : : : : : : : : | | : : : |
Qy 206 ARPRRHAGGPRRLGRCERAMNHSTREAGVPLGLPAPCARRRGGSASLSLPRPRRGAA 255
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 96 ADPA-DPVGHPA-A-P-RAPGEPTRLQPTPRRSGAADPADPVGHPAPRAPGEPT 151
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 266 PEPEPTPVGGSMHAPGRTGSPDRG-FCVSPSPAR-AEATSLLEGALSTRSHSP-SVG 322
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 152 RLQPATPRRSG 162
| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 323 RQHHAGPPTS 333

RESULT 11
ENTRY
TITLE C29149 #type fragment
ORGANISM proline-rich protein - mouse (fragment)
#format_name Mus musculus #common_name house mouse
DATE 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change
20-Mar-1998
ACCESSIONS
C29149
REFERENCE
A92501
#authors
Clements, S.; Mehanasho, H.; Carlson, D.M.
#journal
J. Biol. Chem. (1983) 260:13471-13477
#title
Novel multigene families encoding highly repetitive peptide
sequences. Sequence analyses of rat and mouse proline-rich
protein cDNAs
#cross-references
MUTID:86033799
#contents
Clone pUMPA40
#accession
C29149
##molecule_type
mRNA
##residues
1-227 ##label
CLE
##cross-references
GB:M11902; NID:g200544; PID:g200545
CLASSIFICATION
#superfamily
proline-rich protein
SUMMARY
length 227 #checksum 1378

Query Match 1.8%; Score 125; DB 2; Length 227;
Best Local Similarity 27.3%; Pred. No. 3,67e-02;
Matches 41; Conservative 34; Mismatches 67; Indels 8; Gaps 8;

Db 49 QQGPRPGGQPRRPPGCGGQPRP-PQGPPPGGQPRRPPGCGGQPRP-PQ 106
| : : : | | : : : : : : : : : : : : : : : | | : : : |
Qy 199 QLGATQARPPHAS-GPRRLGCE-RAAMNHSYEDACVGLPLPAGARRGGGSASRLPL 256
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 107 GPRPGGQPRRPPGCGGQPRP-PQPRRPPGCGGQPRRPPGCGQPRRPPAPQPRPPGPR 165
| : : : | | : : : : : : : : : : : : : : : | | : : : |
Qy 257 PKRPRGGAPEPRTPVGGSMHAPGRTGSPDRGFCVSPSPAR-PAEATSLLEGALSGTR 315
: : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 166 PRAGQPRPPQ-GPPPTGQPRPPQGP-PP 193
| : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 316 HSHPSVGRQHAGPPTSRRPRPDTCPP 345

RESULT 12
ENTRY
TITLE D34768 #type complete
ORGANISM ORF4 protein - Orf virus (strain NZ2)
#format_name Orf virus
DATE 23-Aug-1991 #sequence_revision 23-Aug-1991 #text_change

```

	ACCESSIONS	31-Oct-1997
	REFERENCE	D34768
	#authors	A34768
	#journal	Frazer, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.
	#title	Virology (1990) 176:379-389
		Sequence analysis of the inverted terminal repetition in the genome of the parpoxvirus, orf virus.
	#cross-references	MUID:90266454
	#accession	D34768
	#status	Preliminary
	##molecule_type	DNA
	##residues	1-264 ##label FRA
	##cross-references	GB:M30023; EMBL:M37623; NID:g332561; PID:g332566
SUMMARY		#length 264 #molecular-weight 25613 #checksum 1957
Query Match		1.8%; Score 129; DB 2; Length 264;
Best Local Similarity		28.2%; Pred.No.1,2ie-02;
Matches	33; Conservative	40; Mismatches 37; Indels 7; Gaps 5;
Db	150 RAPGAGTTPAPAPAAASAASRHHRRPPARSSSPARPRPAPARAALAPRARSSS-GS	208
Oy	223 RAMNSYR-EGVGLGPAPAPARRGGASAKSLPLRP--R-RGAAPPERPTWQGS	277
Db	209 RASSGPPRSARS-SAASGRPASGPAPAPAPAPASSAFSAGSGAAGCAPAGN	264
Oy	278 WAHPERTGPSDRGCYVSFAPPAEATSTLEGALSGTRHSHPSVGROHAHGPFSTR	334
RESULT	13	
ENTRY		S19560 #type complete
TITLE		proline-rich protein MP4 - mouse
ORGANISM		#format_name Mus musculus #common_name house mouse
DATE		22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 20-Mar-1998
ACCESSIONS		S19560; S22570
REFERENCE		S19560
#authors		Roberts, S.G.E.; Layfield, R.; Bannister, A.J.; McDonald, C.J.
#journal		Eur. J. Biochem. (1991) 202:969-974
#title		Gene sequence of mouse B-type proline-rich protein MP4. Transcriptional start point and an upstream phylogenetic footprint with ets-like and rel/NFkB-like elements.
#cross-references		MUID:9211548
#accession		S19560
#status		Preliminary
##molecule_type		DNA
##residues		1-300 ##label ROB
##cross-references		GB:X58438; NID:g53181; PID:g53182
REFERENCE		S22570
#authors		Roberts, S.G.E.; Layfield, R.; McDonald, C.J.
#journal		Nucleic Acids Res. (1991) 19:5205-5211
#title		The mouse proline-rich protein MP6 promoter binds 12oprenaline-inducible paroid nuclear proteins via a highly conserved NFkB/rel-like site.
#cross-references		MUID:9202026
#accession		S22570
##molecule_type		DNA
##residues		1-14 ##label RO2
##cross-references		EMBL:X61126
CLASSIFICATION		#superfamily proline-rich protein
SUMMARY		#length 300 #molecular-weight 31129 #checksum 8443
Query Match		1.8%; Score 131; DB 2; Length 300;
Best Local Similarity		25.2%; Pred.No.6,89e-03;
Matches	34; Conservative	40; Mismatches 54; Indels 7; Gaps 7;
Db	131 PGNOGPPPOG-QPGQARP-OPNGOQGPPPPGGQQARPORPGNGOGPP-OGCPNP-PPR	166
Oy	202 LAATQARRPHASGPRRRIGCERANNHSVRXGVGLGAPAPARRRGGSASASLSPLPRPR	261
Db	187 PGNOGPPPOGQPORPPPOGNOGPPPOGQ-PAPRRPGNOGPPPOGPPRTGNQ	245
Oy	262 RGAAPEPERTVGGGSAHNRGTGSDRGCVVSAPRAP-AEATSLLEGALSGTRHSHPS	320

DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE ALPHA TRANS-INDUCING PROTEIN (ALPHA-TIP).
 OS BOVINE HERPESVIRUS TYPE 1 (STRAIN PB-2).
 OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; ALPHAHERPESVIRINAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 93012995.
 RA CARPENTER D.E., MISRA V.;
 RL GENE 119:259-263(1992).
 CC -1- FUNCTION: RESPONSIBLE FOR TRANSCRIPTIONAL ACTIVATION OF IMMEDIATE
 CC -1- PROMOTERS (ALPHA GENES) (BY SIMILARITY).
 CC -1- SIMILARITY: TO OTHER HERPESVIRUSES ALPHA TRANS-INDUCING PROTEIN.
 DR EMBL: Z11610.E264419; .
 DR PIR: S24329; S24329.
 DR PIR: JCI306; JCI306.
 KW TRANSCRIPTION REGULATION; TRANS-ACTING FACTOR; DNA-BINDING.
 QO SEQUENCE 504 AA; 54028 MW; 79F42020 CRC32;

Query Match	2.2%;	Score 153;	DB 1;	Length 504;
Best Local Similarity	31.3%;	Pred. No. 1.47e-07;		
Matches	46;	Conservative	33;	Mismatches 56;
			Indels 12;	Gaps 11;

[illegible]

RESULT	3			
ID	PRIM.HUMAN	STANDARD:	PRT:	234 AA.
AC	P10161; P02813;			
DT	01-MAR-1989 (REL. 10, CREATED)			
DT	01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)			
DT	01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)			
DE	SALIVARY PROLINE-RICH PROTEIN PO (ALLELE M) (CONTAINS: PEPTIDE P-D)			
DE	(FRAGMENT).			
CS	PRB4.			
CS	HOMO SAPIENS (HUMAN).			
OC	EDUAROTA; METAFOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;			
OC	EUTHERIA; PRIMATES.			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	LYONS K.M., STEIN J.H., SMITHIES O.;			
RL	GENETICS 120:255-265(1988).			
RN	[2]			
RP	SEQUENCE OF 165-234.			
RX	MEDLINE; 83186122.			
RA	SAITOH E., ISEMURA S., SANADA K.;			
RL	J. BIOCHEM. 93:495-502(1983).			
DR	EMBL; X07704; E265547; -.			
DR	PIR; A03295; PIHUSD.			
DR	PIR; S03175; S03175.			
DR	HSSP; P19999; ICIIG.			
DR	MIM; 168730; -.			
DR	MIM; 180990; -.			
KW	REPEAT; PAROTID GLAND; SALIVA; MULTIGENE FAMILY.			
FT	NON_TER	1		
FT	CHAIN	165	234	PEPTIDE P-D.
SO	SEQUENCE	234 AA;	23676 MW;	ED2D4ADC CRC32;
Query Match		2.1%;	Score 150;	DB 1; Length 234;
Best Local Similarity		29.3%;	Pred. No. 4,48e-07;	
Matches	43; Conservative	41;	Mismatches 53;	Indels 10; Gaps 10;

```

Db      88  GRRP-QGSGNOSQG-PRPRPKGKEPRP-QQG-NQSHRPPRP-CKPR-PPQSGNOSG 144
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      194  GRRPLQJLQAAQABRPAPASGRRRLCCERAMNNSVREAGVGLGLPAPGARRRGSASRS 253
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      142  -P-PRHEKREPRRPOGNSRSASRPDKKQKQPOOEGNKQGPQRPGRKQGPQPGN 199
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      254  LPLRRKRRGAPRRPERTRVYGQSGMAHGRKRGSDR-GFCVVSAPRAE-EATISLEGAL 311
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      200  PQQRPAPRAGKQGPQRPQGGRRPRP 226
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      312  SGTNRSHPSVGRQHNHAGPSTSRPRP 338
      111  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

ID	PRFL	HUMAN	STANDARD;	PRT;	276	AA.
1D	REP.	4				
AC	P10162;	P02813;				
DT	01-MAR-1989	(REL. 10, CREATED)				
DT	01-MAR-1989	(REL. 10, LAST SEQUENCE UPDATE)				
DT	01-OCT-1994	(REL. 30, LAST ANNOTATION UPDATE)				
DE	SALIVARY PROLINE-RICH PROTEIN PO (ALLELE K)	(CONTAINS: PEPTIDE P-D)				
DE	(FRAGMENT).					
GN	PRB4.					
OS	HOMO SAPIENS (HUMAN).					
OC	EUMAROTA: METAZOA: CHORDATA: VERTEBRATA: TETRAPODA: MAMMALIA;					
OC	EUHERETIA: PRIMATES.					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RA	MEDLINE; 89121439.					
RL	LYONS K.M., STEIN J.H., SMITHIES O.;					
RL	GENETICS 120:255-265(1988).					
RN	[2]					
RP	SEQUENCE OF 207-276.					
RP	MEDLINE; 83186122.					
RA	SATOH E., ISEMURA S., SANADA K.;					
RL	J. BIOCHEM. 93:495-502(1983).					
DR	EMBL; X07715; E4806; ALT_SEQ.					
DR	PIR; A03295; PIHUSD.					
DR	PIR; S03176; S03176.					
DR	HSSP; P19999; 1CLG.					
DR	MIM; 168730; -.					
DR	MIM; 180990; -.					
KT	REPEAT: PAROTID GLAND; SALIVA; MULTIGENE FAMILY.					
FT	NON_TER	1				
FT	CHAIN	207	276			PEPTIDE P-D.
SO	SEQUENCE	276	AA; 27816	MM; 4838945A	CRC32;	

[illegible]

Db 157 GVLLETETPRFVSVPQKATGSPFRLPLK-V-VHDDPPHSLRPGSRLOPT 214
Qy 194 GPFLYGLCAQAPRPAS-CPR-RRLGCEBMANHSTRKGVPLGLPAGAR-RR-GGS 249
Db 215 VRRPLAPNCHSPROPPLSDDDGLGAPRLAHSITDPPRRI-TPGSGN-THDLRPL 272
Qy 250 ASRLSLPLKRRKRGGAAPERTPVG-GQSA-HGGRTPGSDKFCVSPAPRAPEKSTSL 307

DI 13-AUG-1987 (REL. 05, CREATED)
DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE PROLINE-RICH PROTEIN MP-3 (FRAGMENT).
GN PRP.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA.
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 86059475.
RA ANN D.K., CARLSON D.M.;
DR J. BIOL. CHEM. 260:15863-15872(1985).
RL EMBL; M12100; G200549; .
DR HSSP; P19999; 1CLG.
KW REPEAT; SALIVA.
FT NON_TER 1 1

```

DE BONE MORPHOGENETIC PROTEIN 8 PRECURSOR (BMP-8) (OSTEOGENIC PROTEIN 2)
DE (OP-2).
GN BMP8.
OS HOMO SAPIENS (HUMAN).
OC EUTHEROTA, METAHOA: CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA, PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-HIPOCAMPUS;
RX MEDLINE; 93094231.
RA OEKAYAKA E., SCHNEGELSBERG P.N.J., JIN D.F., CLIFFORD G.M.,
RA WARREN F.D., DRIER E.A., OPPERMAN H.;
RL J. BIOL. CHEM. 267:25220-25227(1992).
CC -1- FUNCTION: INDUCES CARTILAGE AND BONE FORMATION. MAY BE THE
CC OSTEOINDUCTIVE FACTOR RESPONSIBLE FOR THE PHENOMENON OF
CC EPTHELIAL OSTEOGENESIS. PLAYS A ROLE IN CALCIUM REGULATION
CC AND BONE HOMEOSTASIS (BY SIMILARITY).
CC -1- SUBUNIT: HOMODIMER, DISULFIDE-LINKED.
CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
DR EMBL; M97016; G189390; -.
DR PIR; A45056; A45056.
DR HSSP; P08112; IITFG.
DR MIM; 602284; -.
DR PROSITE; PS00250; TGF_BETA_1.
KW SIGNAL: GROWTH FACTOR; CYTOKINE: BONE; CARTILAGE; GLYCOPROTEIN.
FT SIGNAL 1
FT PROPEP 20 263 POTENTIAL.
FT CHAIN 264 402 BONE MORPHOGENETIC PROTEIN 8.
FT DISULFD 301 367 BY SIMILARITY.
FT DISULFD 330 399 BY SIMILARITY.
FT DISULFD 334 401 BY SIMILARITY.
FT DISULFD 366 366 INTERCHAIN (BY SIMILARITY).
FT CARBOHD 158 158 POTENTIAL.
FT CARBOHD 343 343 POTENTIAL.
SQ SEQUENCE 402 AA; 44764 MW; 8F20C81B CRC32;

Query Match 1.8%; Score 125; DB 1; Length 402;
Best Local Similarity 39.7%; Pred. No. 2,95e-03;
Matches 25; Conservative 12; Mismatches 21; Indels 5; Gaps 5

Db 68 RLP 70
OY 253 SLP 255

RESULT 11
ID MFH1_HUMAN STANDARD; PRT; 501 AA.
AC O99958;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE MESENCHYME FORK PROTEIN 1 (MFH-1 PROTEIN) (TRANSCRIPTION FACTOR
DE FKH-14).
GN FKHL14 OR MFH1.
OS HOMO SAPIENS (HUMAN).
OC EUTHEROTA, METAHOA: CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA, PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97312712.
RA MURA N., IIDA K., KAKIMURA H., YANG X.-L., SUGIYAMA T.;
RL GENOMICS 41:489-492(1997).
CC -1- FUNCTION: MIGHT BE INVOLVED IN THE FORMATION OF SPECIAL
CC MESENCHYMAL TISSUES (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (PROBABLE).
CC -1- SIMILARITY: CONTAINS 1 FORK-HEAD DOMAIN.
DR EMBL; Y08223; E303016; -.
DR MIM; 602402; -.
DR PROSITE; PS00657; FORK_HEAD_1; 1.

```



```

DR EMBL: M97017; G200141; -.
DR MGD: MGI.104515; BMP8A.
DR PROSITE: PS00250; TGF-BETA; 1.
KM SIGNAL; GROWTH FACTOR; CYTOKINE; BONE; CARTILAGE; GLYCOPROTEIN.
FT SIGNAL 1 19 POTENTIAL.
FT PROPEP 20 260 POTENTIAL.
FT CHAIN 261 399 BONE MORPHOGENETIC PROTEIN 8A.
FT DISULFID 298 364 BY SIMILARITY.
FT DISULFID 327 396 BY SIMILARITY.
FT DISULFID 331 398 BY SIMILARITY.
FT DISULFID 363 363 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 155 155 POTENTIAL.
FT CARBOHYD 340 340 POTENTIAL.
SQ SEQUENCE 399 AA; 44764 MW; 7075A1ED CRC32;

Query Match 1.7%; Score 118; DB 1; Length 399;
Best Local Similarity 38.7%; Pred. No. 2.87e-02;
Matches 24; Conservative 12; Mismatches 22; Indels 4; Gaps 4;

Db 12 GLACALGGGHGPRP-HTC-PORRIG-ARERRMOREILLAVLGLPGR-PRPRAQPAAR 67
QY 194 GPPYQLGAATQARPPHAGSPRRRLGCGERAMNSVREAGVPLGLPAPGARRRGGSASRS 253
Db 68 QP 69
QY 254 LP 255

RESULT 15
ID UL61_HCMVA STANDARD; PRT; 431 AA.
AC P16818;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 01-FEB-1991 (REL. 17, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL PROTEIN UL61.
GN UL61.
OS HUMAN CYTOMEGALOVIRUS (STRAIN AD169).
OC VIRIDAE; DS-DNA ENVELOPED VIRUSES; HERPESVIRIDAE; BETAHERPESVIRINAE.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 90269039.
RA CHEE M.S., BANKIER A.T., BECK S., BOHNT R., BROWN C.M., CERNY R.,
RA HORSNELL T., HUTCHISON C.A. III, KOZARIDES T., MARTIGNETTI J.A.,
RA PREDIE E., SATCHELL S.C., TOMLINSON P., WESTON K.M., BARRELL B.G.;
RL CURR. TOP. MICROBIOL. IMMUNOL. 154:125-169(1990).
DR EMBL: X17403; E27280; -.
DR FJR; S09824; S09824.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 431 AA; 44309 MW; 232AB9D7 CRC32;

Query Match 1.7%; Score 118; DB 1; Length 431;
Best Local Similarity 32.1%; Pred. No. 2.87e-02;
Matches 27; Conservative 20; Mismatches 31; Indels 6; Gaps 6;

Db 268 RGGGKPPPLGSPRA-TDGNRDPG-A-CVPAPGRMGSSGSGRGTRGCGPERAAPGARP 324
QY 211 HASGPRRLGCGERAMNSVREAGVPLGLPAPGARRRGGSAS-RSLPLKPRRGAAPPE 269
Db 325 TAPDGAPGRWDGPAAGPAGLGRG 348
QY 270 RTPVG-QGSWAHPGRTGPPS-DRG 291

```

Search completed: Fri Dec 18 18:14:00 1998
Job time : 79 secs.

 WIRE5RECH
 (TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Fri Dec 18 18:37:59 1998; MasPar time 18.72 Seconds
 Tabular output not generated. 800.790 Million cell updates/sec

Title: >US-08-951-733-14
 Description: (640-940) from US08951733.pep (2 of 2)
 Perfect Score: 2214
 Sequence: 1 TSLRFRPKDGLRPIVND.....NLRKYVNPVEDEALGTA 301

Scoring table:
 Gap 11
 PAM 150

Searched: 165420 seqs, 49795644 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

splembl6
 1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mmc 8:sp_organelle
 9:sp_plant 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp_vertebrate 14:sp_virus

Statistics: Mean 47.179; Variance 86.093; scale 0.548

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	2214	100.0	1132	4	014746	TELOMERASE REVERSE TRA	0.00e+00
2	2214	100.0	1132	4	014783	TELOMERASE CATALYTIC S	0.00e+00
3	1596	72.1	1122	11	070372	TELOMERASE REVERSE TRA	0.00e+00
4	375	16.9	988	3	013339	TELOMERASE REVERSE TRA	3.34e-52
5	375	16.9	989	3	013338	TELOMERASE SUBUNIT P12	1.81e-17
6	201	9.1	1031	5	000939	CHROMOSOME XII COSMID	4.10e-08
7	148	6.7	884	3	006163	HYPOTHETICAL 10.7 KD P	6.37e-02
8	108	4.9	94	2	051242	GTP-BINDING PROTEIN (G	6.37e-02
9	108	4.9	366	2	025293	NY3.4.	2.29e-01
10	104	4.7	561	5	045321	NITRITE REDUCTASE (NAD	2.29e-01
11	104	4.7	1002	2	066583	FLAVONOID 3',5'-HYDROX	4.29e-01
12	102	4.6	516	10	096581	POLYPROTEIN PRECURSOR	5.85e-01
13	101	4.6	3410	14	089276	POLYPROTEIN PRECURSOR	5.85e-01
14	101	4.6	3410	14	089275	YAFJ.	1.08e+00
15	99	4.5	280	2	050469	HYPOTHETICAL PROTEIN M	1.08e+00
16	99	4.5	372	1	058694	ORF (FRAGMENT)	1.08e+00
17	99	4.5	380	3	006791	NATURAL-RELATED PROTE	1.08e+00
18	99	4.5	425	2	052231	NEUROFILAMENT-LIKE PRO	7.96e-01
19	100	4.5	557	5	006829	HYPOTHETICAL 66.2 KD P	7.96e-01
20	100	4.5	570	3	014025		

RESULT	ID	PRELIMINARY	PRT	1132 AA	ALIGNMENTS
AC	014746				
AD	014746				
DI	01-JAN-1998 (TREMBLREL. 05, CREATED)				
DT	01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)				
DT	01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)				
DE	TELOMERASE REVERSE TRANSCRIPTASE.				
GN	HTRT.				
OS	HOMO SAPIENS (HUMAN).				
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;				
OC	EUTHERIA; PRIMATES.				
NP	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=KIDNEY;				
RX	MEDLINE: 97400623.				
RA	NARAYANA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,				
RL	LINGNER J., HARLEY C.B., CECH T.R.;				
DR	SCIENCE 277:955-959(1997).				
KW	EMBL: AF015950; G2330017; -				
SO	RNA-DIRECTED DNA POLYMERASE.				
SEQUENCE	1132 AA; 126995 MW; 2DFBEDF3 CRC32;				
Query Match	100.0%; Score 2214; DB 4; Length 1132;				
Best Local Similarity	100.0%; Pred. No. 0.00e+00;				
Matches	301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
DB	618 TSLRFRPKDGLRPIVNDYVVGARTFRREKRAELSRKALFSVYNEKARRPGLG 677				
QY	640 TSLRFRPKDGLRPIVNDYVVGARTFRREKRAELSRKALFSVYNEKARRPGLG 699				
DB	678 ASVLDDIHRAMRFVVRVADPPPELYFYKVVNTGAYDIPDRLTVEYASIRKPN 737				
QY	700 ASVLDDIHRAMRFVVRVADPPPELYFYKVVNTGAYDIPDRLTVEYASIRKPN 759				
DB	738 TYCVARRVAVVQAAAGHVKAFKSHVSTLTDLPYMRQFVAHLOETSPLRDAVVEIOSS 797				
QY	760 TYCVARRVAVVQAAAGHVKAFKSHVSTLTDLPYMRQFVAHLOETSPLRDAVVEIOSS 819				
DB	798 LNEASSGLFDVFLRMCHAVIRKSKSYVOCGIPGSGILSTLSCSLCYGDMENKLFNGI 857				
QY	820 LNEASSGLFDVFLRMCHAVIRKSKSYVOCGIPGSGILSTLSCSLCYGDMENKLFNGI 879				
DB	858 RRDGILLRVDLFDLVTPLHTLPAKTFELTLVGVPEYCYVNLRTVYVNPVEDEALGCT 917				

```
QY 880 RRDGLLRLVDDFLVTPHLTHAKTFLRTLVKGVPEYGCYVNLKRTVVPVEDEALGCT 939
Db 918 A 918
QY 940 A 940

RESULT 2
ID 014783 PRELIMINARY; PRT: 1132 AA.
AC 014783;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TELOMERASE CATALYTIC SUBUNIT.
CN HEST2.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE; 97433088.
RA MEYERSON M., COONTER C.M., EATON E.N., ELLISEN L.M., STEINER P.,
RA CADDE S.D., ZIAGRA L., BEIJERBERGEN R.L., DAVIDOFF M.J., LIU Q.,
RA BACCHETTI S., HABER D.A., WEINBERG R.A.;
RL CELL 90:785-795(1997).
DR EMBL; AF018167; G2347129; -.
SQ SEQUENCE 1132 AA; 126937 MW; C1E5E2AF CRC32;

Query Match 100.0%; Score 2214; DB 4; Length 1132;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 618 TSLRFLTPKPGGLRTVMDYVVGARTFRREKRAERLTSRYKALFSVLYTERARRPGLLG 677
QY 640 TSLRFLTPKPGGLRTVMDYVVGARTFRREKRAERLTSRYKALFSVLYTERARRPGLLG 699
Db 678 ASVLGLDDIHRAMRTFVLVRADQPPPELYFVKVDTGAYDITPODRLEVIASIIKPPON 737
QY 700 ASVLGLDDIHRAMRTFVLVRADQPPPELYFVKVDTGAYDITPODRLEVIASIIKPPON 759
Db 738 TYCVARYAVVOKAHGHVKAFAKFSHVSITLTDLPYMRQFVAHLQETSPLRDAVIEQSS 797
QY 760 TYCVARYAVVOKAHGHVKAFAKFSHVSITLTDLPYMRQFVAHLQETSPLRDAVIEQSS 819
Db 798 LNEASSGFLDFPLRMCCHAVIRKRSYVOCGIPQGSITSLTSLCCTGDMENKLFAGI 857
QY 820 LNEASSGFLDFPLRMCCHAVIRKRSYVOCGIPQGSITSLTSLCCTGDMENKLFAGI 879
Db 858 RRDGLLRLVDDFLVTPHLTHAKTFLRTLVKGVPEYGCYVNLKRTVVPVEDEALGCT 917
QY 880 RRDGLLRLVDDFLVTPHLTHAKTFLRTLVKGVPEYGCYVNLKRTVVPVEDEALGCT 939
Db 918 A 918
QY 940 A 940

RESULT 3
ID 070372 PRELIMINARY; PRT: 1122 AA.
AC 070372;
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RN SEQUENCE FROM N.A.
RX GREENBERG R.A., ALLSOPP R.C., CHIN L., MORIN G.B., DEPINHO R.A.;
RA ONCOGENE 0:0-0(1998).
DR EMBL; AF051911; G3005592; -.
RN RNA-DIRECTED DNA POLYMERASE.
```

```
SO SEQUENCE 1122 AA; 127977 MW; 22207506 CRC32;

Query Match 72.1%; Score 1596; DB 11; Length 1122;
Best Local Similarity 65.6%; Pred. No. 0.00e+00;
Matches 198; Conservative 68; Mismatches 33; Indels 3; Gaps 2;

Db 610 RLRFPPNGILRPVNVNYSWGTGALRRROAOHFTQRLTKTSMUNYERKHPHMLGSS 669
QY 642 RLRFPPNGILRPVNVNYSWGTGALRRROAOHFTQRLTKTSMUNYERKHPHMLGSS 701
Db 670 VLGNADYRTYTRAFVLRVLRADQPPPELYFVKVDTGAYDITPODRLEVIASIIKPPON 739
QY 702 VLGNADYRTYTRAFVLRVLRADQPPPELYFVKVDTGAYDITPODRLEVIASIIKPPON 760
Db 730 YCIRQYAVVRDQSGQVHKSFROVYTLSDLPYMRQFVAHLQETSPLRDAVIEQSS 789
QY 761 YCIRQYAVVRDQSGQVHKSFROVYTLSDLPYMRQFVAHLQETSPLRDAVIEQSS 818
Db 790 SMNSSLTPDFLPHLRHSVYIGDRCTYOCOGIPQGSITSLTSLCCTGDMENKLFAG 849
QY 819 SMNSSLTPDFLPHLRHSVYIGDRCTYOCOGIPQGSITSLTSLCCTGDMENKLFAG 878
Db 850 VORDGLLRLVDDFLVTPHLTHAKTFLRTLVKGVPEYGCYVNLKRTVVPVEDEALGCT 909
QY 879 VORDGLLRLVDDFLVTPHLTHAKTFLRTLVKGVPEYGCYVNLKRTVVPVEDEALGCT 938
Db 910 A 911
QY 939 TA 940

RESULT 4
ID 013339 PRELIMINARY; PRT: 988 AA.
AC 013339;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DE 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE 1.
CN TR1.
OS SCHIZOSACCHAROMYCES POMBE (FUSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCETES.
RN [1]
RN SEQUENCE FROM N.A.
RX STRAIN-972H-.
RX MEDLINE; 97400623.
RA NAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
RA LINGNER J., HARLEY C.B., CECCH T.R.;
RL SCIENCE 277:955-959(1997).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN-972H-.
RA LINE M., RAJANDREAM M.A., BARRELL B.G., VOLCAERT G.;
RL SUBMITTED (MAR-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF015783; G2340168; -.
DR EMBL; AL022299; E1285360; -.
DR PRAM; PR00078; YV.
RN RNA-DIRECTED DNA POLYMERASE.
SQ SEQUENCE 988 AA; 116328 MW; 39C385A7 CRC32;

Query Match 16.9%; Score 375; DB 3; Length 988;
Best Local Similarity 30.7%; Pred. No. 3.34e-52;
Matches 65; Conservative 51; Mismatches 90; Indels 6; Gaps 5;

Db 585 YFVADIRKSCYDRIKQDLMFRIYKRLKDE-FVIRKYATIH-ATSDRAITNVESEAFY 642
QY 729 YFVADIRKSCYDRIKQDLMFRIYKRLKDE-FVIRKYATIH-ATSDRAITNVESEAFY 788
Db 643 FDMVPEEK--VQILSKT--SDILFVDVFDYWKSSSEIFKMLKEHLSGHIVKGNQYL 699
QY 789 TDLPYMRQFVAHLQETSPLRDAVIEQSSLNEASSGFLDFPLRMCCHAVIRKRSYV 848
Db 700 QKVGIPQGSITSLTSLCCTGDMENKLFAGI 759
```

QY 849 QCGGPGSIIITLSCICGDMENKLFAGIRRDG-LLRLVDEFLVTPHLTHAKTELR 907
DB 760 LSLRGEKHNSTLEKTVINPENGIIINT 791
QY 908 TLVRGVPYGCVVNLKRTVNPVEDALGGT 939

RESULT 5
ID 01338 PRELIMINARY; PRT: 989 AA.
AC 01338.
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TELOMERASE REVERSE TRANSCRIPTASE 1.
GN TRT1
OS SCHIZOSACCHAROMYCES POMBE (FISSION YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-972H-;
RX MEDLINE; 97400623.
RA MAKAMURA T.M., MORIN G.B., CHAPMAN K.B., WEINRICH S.L., ANDREWS W.H.,
RA LINGNER J., HARLEY C.B., CECCH T.R.;
RL SCIENCE 277:955-959(1997).
DR EMBL: AF015783; G2340169; -.
DR PFAM: PF00078; rvt.
KM RNA-DIRECTED DNA POLYMERASE.
SQ SEQUENCE 989 AA; 116456 MW; FDE74202 CRC32;

Query Match 16.9%; Score 375; DB 3; Length 989;
Best Local Similarity 30.7%; Pred. No. 3,34e-52;
Matches 65; Conservative 51; Mismatches 90; Indels 6; Gaps 5;

DB 586 YFVRIDKSCYDRKODIMRIVKKLKDP-EVIRKATIH-ATSDRATKNVSEAFSY 643
QY 729 YFVYVDYTGAVDTIPDRLTEVIASIIIPONTYCVRAVAVVQKAHGHVRAKFSHVSTL 788
DB 644 FDMVPEEK--VVOLISMT-SDLEFVVDVDTWTSSEIFMKMKHLSGHVYKIGNSYL 700
QY 789 TDLQYMKQYVAHLOETSPLDVAVIEOSSSLNENSSGLFVFLFKCHHVRIRKGSYV 848
DB 701 QKVGIPGSIISFLCHFYMEDLIDELYSFTKKGSVLLRVVDELFTVKKKDKAKKFLN 760
QY 849 QCGGPGSIIITLSCICGDMENKLFAGIRRDG-LLRLVDEFLVTPHLTHAKTELR 907
DB 761 LSLRGEKHNSTLEKTVINPENGIIINT 792
QY 908 TLVRGVPYGCVVNLKRTVNPVEDALGGT 939

RESULT 6
ID 000939 PRELIMINARY; PRT: 1031 AA.
AC 000939.
DT 01-JUL-1997 (TREMBLREL. 04, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-JAN-1998 (TREMBLREL. 05, LAST ANNOTATION UPDATE)
DE TELOMERASE SUBUNIT P123.
OS EUKARYOTA; MITOCHONDRIAL EUKARYOTES; ALVEOLATA; CILIOPHORA;
OC EUKARYOTAE; MITOCHONDRIAL EUKARYOTES; ALVEOLATA; CILIOPHORA;
OC HYPOTRICHS; EUPLOTIDA; EUPLOTES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97274210.
RA LINGNER J., HUGHES T.R., SHEVCHENKO A., MANN M., LONDEBLAD V.,
RA CECCH T.R.;
RL SCIENCE 276:561-567(1997).
DR EMBL: U95964; G2072336; -.
SQ SEQUENCE 1031 AA; 122562 MW; 21A885CD CRC32;

Query Match 9.1%; Score 201; DB 5; Length 1031;
Best Local Similarity 24.3%; Pred. No. 1,81e-17;
Matches 77; Conservative 82; Mismatches 126; Indels 32; Gaps 21;

DB 512 GKLRILPKKTEFRI--MTF--NKKIVNSDRKTKTLNTRKLSHMLTKLRNMFKDP 567
QY 641 SRLKFIKPGDLRIIVMDYVVGARTERRERKARLTSRKALFS--VLNTERARR-PGL 697
DB 568 FGAVFNIDVDMKKEEYFVCK-WRQVGPKLFEATMDIEKCYDSVNRKLSFTFKTKLL 626
QY 698 LGSVTLGLDDIHRAMRTFVLVRAQDDPPPELYFVKVDTGAVDTIPDRLTEVI-AS-II 755
DB 627 SSDVIMTAQILKRNKNIVYDSKMFRRKEMKDIYRQKFALEGGQYPIFLSVLENDQ 686
QY 756 KPQ--NYCY--RRYAV-Q-KAAGHVRA-KFSHVSTLT-DLQPYMROF-V-AHLQ 804
DB 687 DLNKKTLIVAKRNFKNONLQPIV-IN-ICOVNYINENKGFYKQKGIPOGICVSII 745
QY 805 TSPRLDAVIEOSSSLNENSSGLFVFLFKCHH-VIRKSGSYVQCGGPGSIIITL 863
DB 746 SSFYATLEESSLGFLEDSNPNPNVNLMLRLDYLTLTTOENNAVLEIEKLIIVSR 805
QY 864 CSLCYGDMENK-L-F---AGIRRDG---LLRLVDEFLVTPHLTHAKTELRVLVGV 914
DB 806 ENGRFNNKTLQTSFPL 822
QY 915 EYGCVVNLKRTVNPV 931

RESULT 7
ID 006163 PRELIMINARY; PRT: 884 AA.
AC 006163.
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE CHROMOSOME XII COSMID 8543.
GN LB543.12.
OS SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RX MEDLINE; 97313267.
RA JOHNSTON M., HILLIER L., RILES L., ALBERMANN K., ANDRE B., ANSORGE W.,
RA BERNES V., BRUCKNER M., DELIUS H., DUBOIS E., DUSTERHOFT A.,
RA ENTIAN K.D., FLOETH M., GOFFEAU A., HEBLING U., HEUMANN K.,
RA HEISS-NEITZEL D., HILBERT H., HILGER F., KLEINE K., KOTTER P.,
RA LOUIS E.J., MESSENGY F., MEMES H.W., MIOGA T., MOSTL D.,
RA MULLER-AUER S., NENTWICH U., OBERMAYER B., PIVAVANDI E., POHL T.M.,
RA PORTETELLE D., PURNELLE B., RECHMANN S., RIEGER M., RINKE M., ROSE M.,
RA SCHARFE M., SCHERENS B., SCHOLLER P., SCHWAGER C., SCHWARZ S.,
RA UNDERWOOD A.P., URBESTARAU L.A., VANDENBOL M., VERHAASSEL P.,
RA VIERENDELS F., VOET M., VOLCKAERT G., VOSS H., WAMUTT R., WEDLER E.,
RA WEDLER H., ZIMMERMANN F.K., ZOLLNER A., HANI J., HOHEISEL J.D.;
RL NATURE 387:0-0(0).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA DU Z.;
RL SUBMITTED (MAR-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA WATERSTON R.;
RL SUBMITTED (FEB-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-S288C (AB972);
RA CHERRY J.M.;
RL SUBMITTED (JUL-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U20618; G662136; -.
SQ SEQUENCE 884 AA; 102662 MW; 1A94320F CRC32;

Query Match 6.7%; Score 148; DB 3; Length 884;
Best Local Similarity 26.2%; Pred. No. 4,10e-08;
Matches 68; Conservative 67; Mismatches 93; Indels 32; Gaps 25;

FT CHAIN 210 284 M PROTEIN.
 FT CHAIN 285 777 ENVELOPE PROTEIN.
 FT CHAIN 778 1186 NS1.
 FT CHAIN 1187 1353 POTENTIAL.
 FT CHAIN 1354 1482 POTENTIAL.
 FT CHAIN 1483 2105 NS3.
 FT CHAIN 2107 2393 POTENTIAL.
 FT CHAIN 2394 2505 POTENTIAL.
 FT CHAIN 2506 2506 NS5.
 SQ SEQUENCE 3410 AA; 379301 MW; 06C3D225 CRC32;

Query Match 4.6%; Score 101; DB 14; Length 3410;
 Best Local Similarity 32.3%; Pred. No. 5.85e-01;
 Matches 21; Conservative 14; Mismatches 27; Indels 3; Gaps 3;

Db 1333 GLTPFLG-LCAFLATRFGRSIPVNEALAAAGVYLAGIAPQEMEN-FLGPIAVGGL 1390
 QY 826 GLFDVFLRMCHH-AVRIRGRSYOCGIPQGSILSTLCSLCYGMENKLFAGIRRDGL 884

Db 1391 LMLVY 1395
 QY 885 LRLVY 889

RESULT 15
 ID 050469 PRELIMINARY; PRT; 280 AA.
 AC 050469;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE YAFJ.
 GN YAFJ.
 OS NEISSERIA GONORRHOEA.
 OC PROKARYOTA; GRACILICUTES; SCOTOBACTERIA; AEROBIC RODS AND COCCI;
 CC NEISSERIACEAE.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MS11;
 RA BARTEN R., MEYER T.F.;
 RL SUBMITTED (NOV-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL; AJ002783; E1198593; -;
 SQ SEQUENCE 280 AA; 31556 MW; 34FDDA18 CRC32;

Query Match 4.5%; Score 99; DB 2; Length 280;
 Best Local Similarity 34.3%; Pred. No. 1.08e+00;
 Matches 12; Conservative 10; Mismatches 12; Indels 1; Gaps 1;

Db 64 VRAYQIKSENVVAHIRKASOGOTS-LANTHPFMRE 97
 QY 763 VRAYAVVQKAHGHVKAFAKSHVSTLTDLPYMRQ 797

Search completed: Fri Dec 18 18:39:05 1998
 Job time : 66 secs.

(工)

Release 3.1A John F. Collins, Biocomputing Research Unit
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution platform by Oxford Molecular Ltd

```

Mpsrch_app  protein - protein database search, using Smith-Waterman algorithm
Run on:      Fri Dec 18 18:50:42 1998
Maspar time 22.83 Seconds
672.605 Million cell updates/sec
Tabular output not generated.

```

Accession:	205-08-921-73-24
Description:	(1-949) from US08951733.psp
Percent Score:	949
Sequence:	1 MASGRCVLLKLTNEALNPAI.....PVEDALGCTAVONPAKCL 949
Scoring table:	TABLE uniprotcable

Searched: 131922 segs, 16180660 residues
Post-processing: Minimum Match 0%
Number of clusters: 1000 summaries

```

Outcomes:
  1: part1 2: part2 3: part3 4: part4 5: part5 6: part6 7: part7
  8: part8 9: part9 10: part10 11: part11 12: part12 13: part13
  14: part14 15: part15 16: part16 17: part17 18: part18
  19: part19 20: part20 21: part21 22: part22 23: part23
  24: part24 25: part25 26: part26 27: part27 28: part28
  29: part29

```

Statistics: Mean 3.359; Variance 0.721; scale 4.521

SUMMARY

MON DEC 21 09:20:22 1991

US-00-951-733-14.1a

Page

MON DEC 21 09:28:22 1999

US-08-951-733-14.xm

Page

MON DEC 21 09:28:22 1998

US-08-951-733-14.105

Page 1

Length of match

[illegible]

1354	6	0	5	464	8	R30197	Human	Wild type	gluco	4.7e+02
1355	6	0	5	465	8	R30195	Human	Wild type	gluco	4.7e+02
1356	6	0	5	466	8	R30196	Human	Wild type	gluco	4.7e+02
1357	6	0	5	467	8	R30196	Human	Wild type	gluco	4.7e+02
1358	6	0	5	470	18	R39353	Human	emine tyros	gluco	4.7e+02
1359	6	0	5	471	18	R7262	Human	emine tyros	gluco	4.7e+02
1360	6	0	5	472	18	R7262	Human	emine tyros	gluco	4.7e+02
1361	6	0	5	473	18	R7262	Human	emine tyros	gluco	4.7e+02
1362	6	0	5	474	18	R7262	Human	emine tyros	gluco	4.7e+02
1363	6	0	5	475	28	R4328	Human	metanoma associ	gluco	4.7e+02
1364	6	0	5	476	14	R70099	Human	metanoma associ	gluco	4.7e+02
1365	6	0	5	477	14	R70099	Human	metanoma associ	gluco	4.7e+02
1366	6	0	5	478	14	R70099	Human	metanoma associ	gluco	4.7e+02
1367	6	0	5	480	22	R16922	Human	lipoyasc	endoxin	4.7e+02
1368	6	0	5	480	22	R16922	Human	lipoyasc	endoxin	4.7e+02
1369	6	0	5	480	22	R16922	Human	lipoyasc	endoxin	4.7e+02
1370	6	0	5	480	22	R16922	Human	lipoyasc	endoxin	4.7e+02
1371	6	0	5	481	27	R40812	Human	lipoyasc	endoxin	4.7e+02
1372	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1373	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1374	6	0	5	481	27	R34140	Human	lipoyasc	endoxin	4.7e+02
1375	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1376	6	0	5	481	15	R81346	Human	emine tyros	gluco	4.7e+02
1377	6	0	5	481	15	R81346	Human	emine tyros	gluco	4.7e+02
1378	6	0	5	481	15	R81346	Human	emine tyros	gluco	4.7e+02
1379	6	0	5	481	15	R81346	Human	emine tyros	gluco	4.7e+02
1380	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1381	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1382	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1383	6	0	5	481	15	R80800	Human	lipoyasc	endoxin	4.7e+02
1384	6	0	5	481	15	R80800	Human	lipoyasc	endoxin	4.7e+02
1385	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1386	6	0	5	481	15	R05851	Human	lipoyasc	endoxin	4.7e+02
1387	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1388	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1389	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1390	6	0	5	481	27	R16937	Human	lipoyasc	endoxin	4.7e+02
1391	6	0	5	482	22	R16936	Human	lipoyasc	endoxin	4.7e+02
1392	6	0	5	482	22	R16936	Human	lipoyasc	endoxin	4.7e+02
1393	6	0	5	482	22	R16936	Human	lipoyasc	endoxin	4.7e+02
1394	6	0	5	487	7	R01396	Human	lipoyasc	endoxin	4.7e+02
1395	6	0	5	487	21	R01395	Human	lipoyasc	endoxin	4.7e+02
1396	6	0	5	487	21	R01395	Human	lipoyasc	endoxin	4.7e+02
1397	6	0	5	489	9	R44399	Human	lipoyasc	endoxin	4.7e+02
1398	6	0	5	489	9	R44399	Human	lipoyasc	endoxin	4.7e+02
1399	6	0	5	490	18	R39180	Human	lipoyasc	endoxin	4.7e+02
1400	6	0	5	490	18	R39180	Human	lipoyasc	endoxin	4.7e+02
1401	6	0	5	490	18	R39180	Human	lipoyasc	endoxin	4.7e+02
1402	6	0	5	491	13	R2127	Human	lipoyasc	endoxin	4.7e+02
1403	6	0	5	491	13	R2127	Human	lipoyasc	endoxin	4.7e+02
1404	6	0	5	492	6	R31185	Human	lipoyasc	endoxin	4.7e+02
1405	6	0	5	492	6	R31185	Human	lipoyasc	endoxin	4.7e+02
1406	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1407	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1408	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1409	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1410	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1411	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1412	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1413	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1414	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1415	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1416	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1417	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1418	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1419	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1420	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1421	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1422	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1423	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1424	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1425	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1426	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1427	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1428	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1429	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1430	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1431	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1432	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1433	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1434	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1435	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1436	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1437	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1438	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1439	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1440	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1441	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1442	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1443	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1444	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1445	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1446	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1447	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1448	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1449	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1450	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1451	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1452	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1453	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1454	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1455	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1456	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1457	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1458	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1459	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1460	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1461	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1462	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1463	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1464	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1465	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1466	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1467	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1468	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1469	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1470	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1471	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1472	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1473	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1474	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1475	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1476	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1477	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1478	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1479	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1480	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1481	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1482	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1483	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1484	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1485	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1486	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1487	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1488	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1489	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1490	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	4.7e+02
1491	6	0	5	492	11	R39661	Human	lipoyasc	endoxin	

405	6	0.6	493.1	Q10483	Human prolactin.	4.7e+02
406	6	0.6	493.1	Q10483	Human prolactin.	4.7e+02
407	6	0.6	493.1	Q10483	Human prolactin.	4.7e+02
408	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
409	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
410	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
411	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
412	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
413	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
414	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
415	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
416	6	0.6	497.18	R23914	Human cytochrome P450	4.7e+02
417	6	0.6	503.2	P70139	Sequence of atropine	4.7e+02
418	6	0.6	505.2	P80802	Sequence encoded by	4.7e+02
419	6	0.6	505.2	P80802	Sequence encoded by	4.7e+02
420	6	0.6	508.10	R15196	Human PDI.	4.7e+02
421	6	0.6	508.10	R15196	Human PDI.	4.7e+02
422	6	0.6	508.5	R35397	PDI.	4.7e+02
423	6	0.6	508.1	P10064	Poly(ethylene glycol)	4.7e+02
424	6	0.6	511.1	P10165	Recombinant alpha-amy	4.7e+02
425	6	0.6	511.2	R07574	Alph-amyloid	4.7e+02
426	6	0.6	512.2	R07574	Alph-amyloid	4.7e+02
427	6	0.6	513.6	R12482	Pelz receptor coding	4.7e+02
428	6	0.6	515.23	R14967	PDZ domain	4.7e+02
429	6	0.6	515.23	R14967	PDZ domain	4.7e+02
430	6	0.6	515.5	R35198	NSP-ORC2 immunost	4.7e+02
431	6	0.6	516.1	P80808	NSP-ORC2 fusion prot	4.7e+02
432	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
433	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
434	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
435	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
436	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
437	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
438	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
439	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
440	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
441	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
442	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
443	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
444	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
445	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
446	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
447	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
448	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
449	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
450	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
451	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
452	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
453	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
454	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02
455	6	0.6	521.27	P12884	Sequence of gap prote	4.7e+02

455	6	0	0	564.13	R69133	Cyclic-GMP stimulated
456	6	0	0	564.20	W1157	Heart calcium/calmodu
457	6	0	0	564.23	R20812	Protein of cDNA, liver
458	6	0	0	564.25	R20812	Protein of cDNA, liver
459	6	0	0	565.22	R20945	H. pylori inner memb
460	6	0	0	565.22	R20945	H. pylori inner memb
461	6	0	0	569.15	R94923	Spam3 3-oxoallopate
462	6	0	0	569.16	R86406	Human matrix metallo
463	6	0	0	569.16	R86406	Human matrix metallo
464	6	0	0	580.25	W43231	Myoblast-derived musc
465	6	0	0	580.25	W43231	Myoblast-derived musc
466	6	0	0	580.22	W25570	hNtr sequence encode
467	6	0	0	580.7	R18151	Acetylcholinesterase
468	6	0	0	580.73	W12362	Acetylcholinesterase
469	6	0	0	582.16	R8607	Human matrix metallo
470	6	0	0	585.7	R16533	988-97 cDNA, 18 prod
471	6	0	0	586.27	W273700	Bacillus sp. alpha-91
472	6	0	0	586.27	W273700	Bacillus sp. alpha-91
473	6	0	0	590.11	R58616	Human c-myc fct uptr
474	6	0	0	592.5	R26420	Human c-myc fct uptr
475	6	0	0	598.19	R98344	Cyclooctatetraene
476	6	0	0	602.7	W13680	CSF/transformed human p
477	6	0	0	602.7	W13680	Human prolymph. mena
478	6	0	0	602.7	R120909	Human prolymph. mena
479	6	0	0	614.5	R28112	Protd. cDNA, tissue of
480	6	0	0	614.5	W11156	Riboprotein calcium/c
481	6	0	0	614.11	R69312	Cyclic-GMP stimulated
482	6	0	0	614.11	R69312	Cyclic-GMP stimulated
483	6	0	0	613.11	R80814	Human c-myc fct uptr
484	6	0	0	613.11	R80814	Human c-myc fct uptr
485	6	0	0	644.11	R12811	anti poly(peptide), and
486	6	0	0	645.6	R12811	anti poly(peptide), and
487	6	0	0	657.23	W00553	H. pylori transducer
488	6	0	0	613.19	R68143	GS/ human p57 tumor
489	6	0	0	667.5	R25510	Endothelial tissue cr
490	6	0	0	667.5	R25510	Endothelial tissue cr
491	6	0	0	684.21	W12157	Drosophila frizzled-2
492	6	0	0	697.37	W12324	Staphylococcus aureus
493	6	0	0	700.2	R11354	Cellular Receptor 2.1
494	6	0	0	704.11	R58176	Human binding protein
495	6	0	0	704.11	R58176	Human binding protein
496	6	0	0	704.17	R58577	MD-40 domain-con t
497	6	0	0	704.20	W06079	Drosophila 74Vb1-shd1
498	6	0	0	707.2	R07077	97-RD human c-myc
499	6	0	0	710.28	W13145	Heat-resistant beta-g
500	6	0	0	710.28	W13145	Heat-resistant beta-g
501	6	0	0	711.4	R10166	Human b1rg receptor
502	6	0	0	711.4	R10166	Chicken basic fibrobl
503	6	0	0	719.17	R14495	Human b1rg receptor
504	6	0	0	719.17	R14495	Human b1rg receptor
505	6	0	0	759.23	W07709	Human p33 protein, A
506	6	0	0	759.23	W07709	Human p33 protein, A

[illegible]

558	6	0	6	1063	4	RJ3916	Shubella vitru strain
559	6	0	6	1068	9	RJ3441	P110
560	6	0	6	1068	28	RK639	3-X-musae J10 K
561	6	0	6	1068	3	RJ3916	Shubella vitru strain
562	6	0	6	1085	3	PL046	A. nidulans phospholipid
563	6	0	6	1091	6	RJ3426	Translated from 5' re
564	6	0	6	1091	6	RJ3426	Translated from 5' re
565	6	0	6	1091	6	RJ3426	Alpha 66 integrin sub
566	6	0	6	1104	17	RJ4957	NR 4X DNA-binding pro
567	6	0	6	1104	17	RJ4957	NR 4X DNA-binding pro
568	6	0	6	1113	11	RJ3133	Unpurified suppressor
569	6	0	6	1117	16	RJ0076	Tumour suppressor, pro
570	6	0	6	1116	5	RJ3136	Tumour suppressor, pro
571	6	0	6	1182	3	RJ5157	Abelson Related Gene,
572	6	0	6	1182	3	RJ5157	Abelson Related Gene,
573	6	0	6	1187	28	WJ4290	Alu retrotransposon
574	6	0	6	1188	28	WJ4291	Alu retrotransposon
575	6	0	6	1230	21	WJ4561	Alu retrotransposon
576	6	0	6	1245	21	WJ3872	Alu retrotransposon
577	6	0	6	1245	21	WJ3872	Bacillus thuringiensis
578	6	0	6	1257	9	RJ6627	Neurospora crassa
579	6	0	6	1257	9	RJ6627	Neurospora crassa
580	6	0	6	1311	11	RJ8823	Neurospora crassa
581	6	0	6	1311	11	RJ8823	Neurospora crassa
582	6	0	6	1315	16	WJ7659	Hypocotyl inducing pr
583	6	0	6	1315	16	WJ7659	Hypocotyl inducing pr
584	6	0	6	1431	14	RJ7405	Rat death associated
585	6	0	6	1436	14	RJ7419	Rat hemoglobin of huma
586	6	0	6	1485	14	WJ7887	Photoreducible lumines
587	6	0	6	1488	14	WJ7887	Photoreducible lumines
588	6	0	6	1488	14	WJ7887	Rat albumin
589	6	0	6	1502	14	RJ7088	Heat-stable supernatant
590	6	0	6	1502	14	RJ7088	Rat albumin
591	6	0	6	1512	14	WJ4523	Rat albumin
592	6	0	6	1512	14	WJ4523	Rat albumin
593	6	0	6	1635	27	WJ4624	Rat albumin
594	6	0	6	1635	27	WJ4624	Rat albumin
595	6	0	6	1657	27	WJ4629	Rat albumin
596	6	0	6	1651	27	WJ4625	Rat albumin
597	6	0	6	1651	27	WJ4625	Rat albumin
598	6	0	6	1653	27	WJ4608	Rat albumin
599	6	0	6	1653	27	WJ4608	Rat albumin
600	6	0	6	1653	27	WJ4608	Rat albumin
601	6	0	6	1653	27	WJ4608	Rat albumin
602	6	0	6	1653	27	WJ4608	Rat albumin
603	6	0	6	1653	27	WJ4608	Rat albumin
604	6	0	6	1653	27	WJ4616	Rat albumin
605	6	0	6	1653	27	WJ4616	Rat albumin
606	6	0	6	1653	27	WJ4616	Rat albumin
607	6	0	6	1653	27	WJ4609	Rat albumin
608	6	0	6	1653	27	WJ4612	Rat albumin

[illegible]

560	5	0	5	1	Pt2808	Von Willbrand factor	4.2e+00
561	5	0	5	1	Pt2809	Von Willbrand factor	4.2e+00
562	5	0	5	1	Pt2810	Von Willbrand factor	4.2e+00
563	5	0	5	1	Pt2811	Immunosuppressive ppv	4.2e+00
564	5	0	5	1	Pt2812	Immunosuppressive ppv	4.2e+00
565	5	0	5	20	R58930	H-deleted variant of	4.2e+00
566	5	0	5	20	R58931	H857-LV6676 of huma	4.2e+00
567	5	0	5	20	R58932	H857-LV6676 of huma	4.2e+00
568	5	0	5	21	R58933	SLF1-L transpolar	4.2e+00
569	5	0	5	21	R58934	SLF1-L transpolar	4.2e+00
570	5	0	5	24	R13100	GPV tandem Leu-rich r	4.2e+00
571	5	0	5	24	R13101	GPV tandem Leu-rich r	4.2e+00
572	5	0	5	24	R13102	Sequence of new enzy	4.2e+00
573	5	0	5	29	R13564	MD-40 domain contig	4.2e+00
574	5	0	5	29	R13565	MD-40 domain contig	4.2e+00
575	5	0	5	35	R0379	Stomach cancer SCM fa	4.2e+00
576	5	0	5	35	R0379	Stomach cancer SCM fa	4.2e+00
577	5	0	5	38	R11	Protein binding compo	4.2e+00
578	5	0	5	41	R8717	Protein binding compo	4.2e+00
579	5	0	5	41	R8717	Protein binding compo	4.2e+00
580	5	0	5	49	R20264	Sequence of parent hu	4.2e+00
581	5	0	5	49	R20264	Sequence of parent hu	4.2e+00
582	5	0	5	54	R22772	Sequence of parent hu	4.2e+00
583	5	0	5	54	R22772	Sequence of parent hu	4.2e+00
584	5	0	5	55	R15166	Human gene clone prod	4.2e+00
585	5	0	5	55	R15166	Human gene clone prod	4.2e+00
586	5	0	5	63	R13908	Branch-chain alpha	4.2e+00
587	5	0	5	63	R13908	Branch-chain alpha	4.2e+00
588	5	0	5	67	R13937	Sequence of human alc	4.2e+00
589	5	0	5	69	Pt1428	GAP fusion protein en	4.2e+00
590	5	0	5	82	R14076	Cytochrome and neuroin	4.2e+00
591	5	0	5	82	R14076	Cytochrome and neuroin	4.2e+00
592	5	0	5	89	R08398	Substituted enzy do	4.2e+00
593	5	0	5	102	R13	Gr-beta2/HRP-2, alpha	4.2e+00
594	5	0	5	107	R23034	Human GTP beta cyclol	4.2e+00
595	5	0	5	111	R78342	Human GTP beta cyclol	4.2e+00
596	5	0	5	111	R78342	Human GTP beta cyclol	4.2e+00
597	5	0	5	112	R05915	Human TLI1 mutator for	4.2e+00
598	5	0	5	112	R05915	Human TLI1 mutator for	4.2e+00
599	5	0	5	113	R05918	PHOS570 (hcc-15-15th	4.2e+00
600	5	0	5	116	R2559	Sequence of mutant pe	4.2e+00
601	5	0	5	116	R2559	Sequence of mutant pe	4.2e+00
602	5	0	5	118	R54158	Mutant NMR1A antigen	4.2e+00
603	5	0	5	118	R54158	Mutant NMR1A antigen	4.2e+00
604	5	0	5	118	R54158	Humanized IgM01 heavy	4.2e+00
605	5	0	5	118	R54158	Humanized IgM01 heavy	4.2e+00
606	5	0	5	120	R09316	Chlaemic protein gag-	4.2e+00
607	5	0	5	120	R09316	Chlaemic protein gag-	4.2e+00
608	5	0	5	123	R5187	Sequence of mutant pe	4.2e+00
609	5	0	5	123	R5187	Sequence of mutant pe	4.2e+00
610	5	0	5	123	R5187	PHOS572 (hcc-15-11)	4.2e+00
611	5	0	5	123	R5187	PHOS572 (hcc-15-11)	4.2e+00
612	5	0	5	123	R5187	Heavy chain variable	4.2e+00
613	5	0	5	128	Pt1318	Sequence of pro-auria	4.2e+00
614	5	0	5	128	Pt1318	Sequence of pro-auria	4.2e+00
615	5	0	5	131	R1176	Human lymphocyte ant	4.2e+00
616	5	0	5	131	R1		4.2e+00
617	5	0	5	131	R1		4.2e+00
618	5	0	5	131	R1		4.2e+00
619	5	0	5	131	R1		4.2e+00
620	5	0	5	131	R1		4.2e+00
621	5	0	5	131	R1		4.2e+00
622	5	0	5	131	R1		4.2e+00
623	5	0	5	131	R1		4.2e+00
624	5	0	5	131	R1		4.2e+00
625	5	0	5	131	R1		4.2e+00
626	5	0	5	131	R1		4.2e+00
627	5	0	5	131	R1		4.2e+00
628	5	0	5	131	R1		4.2e+00
629	5	0	5	131	R1		4.2e+00
630	5	0	5	131	R1		4.2e+00
631	5	0	5	131	R1		4.2e+00
632	5	0	5	131	R1		4.2e+00
633	5	0	5	131	R1		4.2e+00
634	5	0	5	131	R1		4.2e+00
635	5	0	5	131	R1		4.2e+00
636	5	0	5	131	R1		4.2e+00
637	5	0	5	131	R1		4.2e+00
638	5	0	5	131	R1		4.2e+00
639	5	0	5	131	R1		4.2e+00
640	5	0	5	131	R1		4.2e+00
641	5	0	5	131	R1		4.2e+00
642	5	0	5	131	R1		4.2e+00
643	5	0	5	131	R1		4.2e+00
644	5	0	5	131	R1		4.2e+00
645	5	0	5	131	R1		4.2e+00
646	5	0	5	131	R1		4.2e+00
647	5	0	5	131	R1		4.2e+00
648	5	0	5	131	R1		4.2e+00
649	5	0	5	131	R1		4.2e+00
650	5	0	5	131	R1		4.2e+00
651	5	0	5	131	R1		4.2e+00
652	5	0	5	131	R1		4.2e+00
653	5	0	5	131	R1		4.2e+00
654	5	0	5	131	R1		4.2e+00
655	5	0	5	131	R1		4.2e+00
656	5	0	5	131	R1		4.2e+00
657	5	0	5	131	R1		4.2e+00
658	5	0	5	131	R1		4.2e+00
659	5	0	5	131	R1		4.2e+00
660	5	0	5	131	R1		4.2e+00
661	5	0	5	131	R1		4.2e+00
662	5	0	5	131	R1		4.2e+00
663	5	0	5	131	R1		4.2e+00
664	5	0	5	131	R1		4.2e+00
665	5	0	5	131	R1		4.2e+00
666	5	0	5	131	R1		4.2e+00
667	5	0	5	131	R1		4.2e+00
668	5	0	5	131	R1		4.2e+00
669	5	0	5	131	R1		4.2e+00
670	5	0	5	131	R1		4.2e+00
671	5	0	5	131	R1		4.2e+00
672	5	0	5	131	R1		4.2e+00
673	5	0	5	131	R1		4.2e+00
674	5	0	5	131	R1		4.2e+00
675	5	0	5	131	R1		4.2e+00
676	5	0	5	131	R1		4.2e+00
677	5	0	5	131	R1		4.2e+00
678	5	0	5	131	R1		4.2e+00
679	5	0	5	131	R1		4.2e+00
680	5	0	5	131	R1		4.2e+00
681	5	0	5	131	R1		4.2e+00
682	5	0	5	131	R1		4.2e+00
683	5	0	5	131	R1		4.2e+00
684	5	0	5	131	R1		4.2e+00
685	5	0	5	131	R1		4.2e+00
686	5	0	5	131	R1		4.2e+00
687	5	0	5	131	R1		4.2e+00
688	5	0	5	131	R1		4.2e+00
689	5	0	5	131	R1		4.2e+00
690	5	0	5	131	R1		4.2e+00
691	5	0	5	131	R1		4.2e+00
692	5	0	5	131	R1		4.2e+00
693	5	0	5	131	R1		4.2e+00
694	5	0	5	131	R1		4.2e+00
695	5	0	5	131	R1		4.2e+00
696	5	0	5	131	R1		4.2e+00
697	5	0	5	131	R1		4.2e+00
698	5	0	5	131	R1		4.2e+00
699	5	0	5	131	R1		4.2e+00
700	5	0	5	131	R1		4.2e+00
701	5	0	5	131	R1		4.2e+00
702	5	0	5	131	R1		4.2e+00
703	5	0	5	131	R1		4.2e+00
704	5	0	5	131	R1		4.2e+00
705	5	0	5	131	R1		4.2e+00
706	5	0	5	131	R1		4.2e+00
707	5	0	5	131	R1		4.2e+00
708	5	0	5	131	R1		4.2e+00
709	5	0	5	131	R1		4.2e+00
710	5	0	5	131	R1		4.2e+00
711	5	0	5	131	R1		4.2e+00
712	5	0	5	131	R1		4.2e+00
713	5	0	5	131	R1		4.2e+00
714	5	0	5	131	R1		4.2e+00
715	5	0	5	131	R1		4.2e+00
716	5	0	5	131	R1		4.2e+00
717	5	0	5	131	R1		4.2e+00
718	5	0	5	131	R1		4.2e+00
719	5	0	5	131	R1		4.2e+00
720	5	0	5	131	R1		4.2e+00
721	5	0	5	131	R1		4.2e+00
722	5	0	5	131	R1		4.2e+00
723	5	0	5	131	R1		4.2e+00
724	5	0	5	131	R1		4.2e+00
725	5	0	5	131	R1		4.2e+00
726	5	0	5	131	R1		4.2e+00
727	5	0	5	131	R1		4.2e+00
728	5	0	5	131	R1		4.2e+00
729	5	0	5	131	R1		4.2e+00
730	5	0	5	131	R1		4.2e+00
731	5	0	5	131	R1		4.2e+00
732	5	0	5	131	R1		4.2e+00
733	5	0	5	131	R1		4.2e+00
734	5	0	5	131	R1		4.2e+00
735	5	0	5	131	R1		4.2e+00
736	5	0	5	131	R1		4.2e+00
737	5	0	5	131	R1		4.2e+00
738	5	0	5	131	R1		4.2e+00
739	5	0	5	131	R1		4.2e+00
740	5	0	5	131	R1		4.2e+00
741	5	0	5	131	R1		4.2e+00
742	5	0	5	131	R1		4.2e+00
743	5	0	5	131	R1		4.2e+00
744	5	0	5	131	R1		4.2e+00
745	5	0	5	131	R1		4.2e+00
746	5	0	5	131	R1		4.2e+00
747	5	0	5	131	R1		4.2e+00
748	5	0	5	131	R1		4.2e+00
749	5	0	5	131	R1		4.2e+00
750	5	0	5	131	R1		4.2e+00
751	5	0	5	131	R1		4.2e+00
752	5	0	5	131	R1		4.2e+00
753	5	0	5	131	R1		4.2e+00
754	5	0	5	131	R1		4.2e+00
755	5	0	5	131	R1		4.2e+00
756	5	0	5	131	R1		4.2e+00

161	5	0	5	251	8	R42621	Truncated tissue factor
160	5	0	5	248	1	R40521	Antigen (GX)252 (E)47c
159	5	0	5	240	6	R40320	Asn-107 cholera toxin
158	5	0	5	236	6	R40322	Cyt dL2/3
157	5	0	5	236	6	R40322	Mouse epimorphin f2eg
156	5	0	5	235	5	R40324	Mouse epimorphin f2eg
155	5	0	5	235	5	R43553	Mouse epimorphin f2eg
154	5	0	5	235	5	R43553	Mouse epimorphin f2eg
153	5	0	5	235	5	R43553	Mouse epimorphin f2eg
152	5	0	5	235	5	R43553	Mouse epimorphin f2eg
151	5	0	5	235	5	R43553	Mouse epimorphin f2eg
150	5	0	5	235	5	R43553	Mouse epimorphin f2eg
149	5	0	5	235	5	R43553	Mouse epimorphin f2eg
148	5	0	5	235	5	R43553	Mouse epimorphin f2eg
147	5	0	5	235	5	R43553	Mouse epimorphin f2eg
146	5	0	5	235	5	R43553	Mouse epimorphin f2eg
145	5	0	5	235	5	R43553	Mouse epimorphin f2eg
144	5	0	5	235	5	R43553	Mouse epimorphin f2eg
143	5	0	5	235	5	R43553	Mouse epimorphin f2eg
142	5	0	5	235	5	R43553	Mouse epimorphin f2eg
141	5	0	5	235	5	R43553	Mouse epimorphin f2eg
140	5	0	5	235	5	R43553	Mouse epimorphin f2eg
139	5	0	5	235	5	R43553	Mouse epimorphin f2eg
138	5	0	5	235	5	R43553	Mouse epimorphin f2eg
137	5	0	5	235	5	R43553	Mouse epimorphin f2eg
136	5	0	5	235	5	R43553	Mouse epimorphin f2eg
135	5	0	5	235	5	R43553	Mouse epimorphin f2eg
134	5	0	5	235	5	R43553	Mouse epimorphin f2eg
133	5	0	5	235	5	R43553	Mouse epimorphin f2eg
132	5	0	5	235	5	R43553	Mouse epimorphin f2eg
131	5	0	5	235	5	R43553	Mouse epimorphin f2eg
130	5	0	5	235	5	R43553	Mouse epimorphin f2eg
129	5	0	5	235	5	R43553	Mouse epimorphin f2eg
128	5	0	5	235	5	R43553	Mouse epimorphin f2eg
127	5	0	5	235	5	R43553	Mouse epimorphin f2eg
126	5	0	5	235	5	R43553	Mouse epimorphin f2eg
125	5	0	5	235	5	R43553	Mouse epimorphin f2eg
124	5	0	5	235	5	R43553	Mouse epimorphin f2eg
123	5	0	5	235	5	R43553	Mouse epimorphin f2eg
122	5	0	5	235	5	R43553	Mouse epimorphin f2eg
121	5	0	5	235	5	R43553	Mouse epimorphin f2eg
120	5	0	5	235	5	R43553	Mouse epimorphin f2eg
119	5	0	5	235	5	R43553	Mouse epimorphin f2eg
118	5	0	5	235	5	R43553	Mouse epimorphin f2eg
117	5	0	5	235	5	R43553	Mouse epimorphin f2eg
116	5	0	5	235	5	R43553	Mouse epimorphin f2eg
115	5	0	5	235	5	R43553	Mouse epimorphin f2eg
114	5	0	5	235	5	R43553	Mouse epimorphin f2eg
113	5	0	5	235	5	R43553	Mouse epimorphin f2eg
112	5	0	5	235	5	R43553	Mouse epimorphin f2eg
111	5	0	5	235	5	R43553	Mouse epimorphin f2eg
110	5	0	5	235	5	R43553	Mouse epimorphin f2eg
109	5	0	5	235	5	R43553	Mouse epimorphin f2eg
108	5	0	5	235	5	R43553	Mouse epimorphin f2eg
107	5	0	5	235	5	R43553	Mouse epimorphin f2eg
106	5	0	5	235	5	R43553	Mouse epimorphin f2eg
105	5	0	5	235	5	R43553	Mouse epimorphin f2eg
104	5	0	5	235	5	R43553	Mouse epimorphin f2eg
103	5	0	5	235	5	R43553	Mouse epimorphin f2eg
102	5	0	5	235	5	R43553	Mouse epimorphin f2eg
101	5	0	5	235	5	R43553	Mouse epimorphin f2eg
100	5	0	5	235	5	R43553	Mouse epimorphin f2eg
99	5	0	5	235	5	R43553	Mouse epimorphin f2eg
98	5	0	5	235	5	R43553	Mouse epimorphin f2eg
97	5	0	5	235	5	R43553	Mouse epimorphin f2eg
96	5	0	5	235	5	R43553	Mouse epimorphin f2eg
95	5	0	5	235	5	R43553	Mouse epimorphin f2eg
94	5	0	5	235	5	R43553	Mouse epimorphin f2eg
93	5	0	5	235	5	R43553	Mouse epimorphin f2eg
92	5	0	5	235	5	R43553	Mouse epimorphin f2eg
91	5	0	5	235	5	R43553	Mouse epimorphin f2eg
90	5	0	5	235	5	R43553	Mouse epimorphin f2eg
89	5	0	5	235	5	R43553	Mouse epimorphin f2eg
88	5	0	5	235	5	R43553	Mouse epimorphin f2eg
87	5	0	5	235	5	R43553	Mouse epimorphin f2eg
86	5	0	5	235	5	R43553	Mouse epimorphin f2eg
85	5	0	5	235	5	R43553	Mouse epimorphin f2eg
84	5	0	5	235	5	R43553	Mouse epimorphin f2eg
83	5	0	5	235	5	R43553	Mouse epimorphin f2eg
82	5	0	5	235	5	R43553	Mouse epimorphin f2eg
81	5	0	5	235	5	R43553	Mouse epimorphin f2eg
80	5	0	5	235	5	R43553	Mouse epimorphin f2eg
79	5	0	5	235	5	R43553	Mouse epimorphin f2eg
78	5	0	5	235	5	R43553	Mouse epimorphin f2eg
77	5	0	5	235	5	R43553	Mouse epimorphin f2eg
76	5	0	5	235	5	R43553	Mouse epimorphin f2eg
75	5	0	5	235	5	R43553	Mouse epimorphin f2eg
74	5	0	5	235	5	R43553	Mouse epimorphin f2eg
73	5	0	5	235	5	R43553	Mouse epimorphin f2eg
72	5	0	5	235	5	R43553	Mouse epimorphin f2eg
71	5	0	5	235	5	R43553	Mouse epimorphin f2eg
70	5	0	5	235	5	R43553	Mouse epimorphin f2eg
69	5	0	5	235	5	R43553	Mouse epimorphin f2eg
68	5	0	5	235	5	R43553	Mouse epimorphin f2eg
67	5	0	5	235	5	R43553	Mouse epimorphin f2eg
66	5	0	5	235	5	R43553	Mouse epimorphin f2eg
65	5	0	5	235	5	R43553	Mouse epimorphin f2eg
64	5	0	5	235	5	R43553	Mouse epimorphin f2eg
63	5	0	5	235	5	R43553	Mouse epimorphin f2eg
62	5	0	5	235	5	R43553	Mouse epimorphin f2eg
61	5	0	5	235	5	R43553	Mouse epimorphin f2eg
60	5	0	5	235	5	R43553	Mouse epimorphin f2eg
59	5	0	5	235	5	R43553	Mouse epimorphin f2eg
58	5	0	5	235	5	R43553	Mouse epimorphin f2eg
57	5	0	5	235	5	R43553	Mouse epimorphin f2eg
56	5	0	5	235	5	R43553	Mouse epimorphin f2eg
55	5	0	5	235	5	R43553	Mouse epimorphin f2eg
54	5	0	5	235	5	R43553	Mouse epimorphin f2eg
53	5	0	5	235	5	R43553	Mouse epimorphin f2eg
52	5	0	5	235	5	R43553	Mouse epimorphin f2eg
51	5	0	5	235	5	R43553	Mouse epimorphin f2eg
50	5	0	5	235	5	R43553	Mouse epimorphin f2eg
49	5	0	5	235	5	R43553	Mouse epimorphin f2eg
48	5	0	5	235	5	R43553	Mouse epimorphin f2eg
47	5	0	5	235	5	R43553	Mouse epimorphin f2eg
46	5	0	5	235	5	R43553	Mouse epimorphin f2eg
45	5	0	5	235	5	R43553	Mouse epimorphin f2eg
44	5	0	5	235	5	R43553	Mouse epimorphin f2eg
43	5	0	5	235	5	R43553	Mouse epimorphin f2eg
42	5	0	5	235	5	R43553	Mouse epimorphin f2eg
41	5	0	5	235	5	R43553	Mouse epimorphin f2eg
40	5	0	5	235	5	R43553	Mouse epimorphin f2eg
39	5	0	5	235	5	R43553	Mouse epimorphin f2eg
38	5	0	5	235	5	R43553	Mouse epimorphin f2eg
37	5	0	5	235	5	R43553	Mouse epimorphin f2eg
36	5	0	5	235	5	R43553	Mouse epimorphin f2eg
35	5	0	5	235	5	R43553	Mouse epimorphin f2eg
34	5	0	5	235	5	R43553	Mouse epimorphin f2eg
33	5	0	5	235	5	R43553	Mouse epimorphin f2eg
32	5	0	5	235	5	R43553	Mouse epimorphin f2eg
31	5	0	5	235	5	R43553	Mouse epimorphin f2eg
30	5	0	5	235	5	R43553	Mouse epimorphin f2eg
29	5	0	5	235	5	R43553	Mouse epimorphin f2eg
28	5	0	5	235	5	R43553	Mouse epimorphin f2eg
27	5	0	5	235	5	R43553	Mouse epimorphin f2eg
26	5	0	5	235	5	R43553	Mouse epimorphin f2eg
25	5	0	5	235	5	R43553	Mouse epimorphin f2eg
24	5	0	5	235	5	R43553	Mouse epimorphin f2eg
23	5	0	5	235	5	R43553	Mouse epimorphin f2eg
22	5	0	5	235	5	R43553	Mouse epimorphin f2eg
21	5	0	5	235	5	R43553	Mouse epimorphin f2eg
20	5	0	5	235	5	R43553	Mouse epimorphin f2eg
19	5	0	5	235	5	R43553	Mouse epimorphin f2eg
18	5	0	5	235	5	R43553	Mouse epimorphin f2eg
17	5	0	5	235	5	R43553	Mouse epimorphin f2eg
16	5	0	5	235	5	R43553	Mouse epimorphin f2eg
15	5	0	5	235	5	R43553	Mouse epimorphin f2eg
14	5	0	5	235	5	R43553	Mouse epimorphin f2eg
13	5	0	5	235	5	R43553	Mouse epimorphin f2eg
12	5	0	5	235	5	R43553	Mouse epimorphin f2eg
11	5	0	5	235	5	R43553	Mouse epimorphin f2eg
10	5	0	5	235	5	R43553	Mouse epimorphin f2eg
9	5	0	5	235	5	R43553	Mouse epimorphin f2eg
8	5	0	5	235	5	R43553	Mouse epimorphin f2eg
7	5	0	5	235	5	R43553	Mouse epimorphin f2eg
6	5	0	5	235	5	R43553	Mouse epimorphin f2eg
5	5	0	5	235	5	R43553	Mouse epimorphin f2eg
4	5	0	5	235	5	R43553	Mouse epimorphin f2eg
3	5	0	5	235	5	R43553	Mouse epimorphin f2eg
2	5	0	5	235	5	R43553	Mouse epimorphin f2eg
1	5	0	5	235	5	R43553	Mouse epimorphin f2eg

[illegible][illegible]

914	5	0	5	593	R18673	Granulin sequence.	4.21e+007
913	5	0	5	588	15		
912	5	0	5	587	R11000	Human beta2-microglobulin cDNA	4.21e+007
911	5	0	5	582	R11000	Human beta2-microglobulin cDNA	4.21e+007
910	5	0	5	577	R14113	Malino acid sequence of	4.21e+007
909	5	0	5	568	R15366	Human HSP70 protein.	4.21e+007
908	5	0	5	567	R15362	Mouse HSP70 protein.	4.21e+007
907	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
906	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
905	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
904	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
903	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
902	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
901	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
900	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
899	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
898	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
897	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
896	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
895	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
894	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
893	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
892	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
891	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
890	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
889	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
888	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
887	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
886	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
885	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
884	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
883	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
882	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
881	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
880	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
879	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
878	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
877	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
876	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
875	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
874	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
873	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
872	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
871	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
870	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
869	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
868	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
867	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
866	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
865	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
864	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
863	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
862	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
861	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
860	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
859	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
858	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
857	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
856	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
855	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
854	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
853	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
852	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
851	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
850	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
849	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
848	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
847	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
846	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
845	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
844	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
843	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
842	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
841	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
840	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
839	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
838	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
837	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
836	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
835	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
834	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
833	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
832	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
831	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
830	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
829	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
828	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
827	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
826	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
825	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
824	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
823	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
822	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
821	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
820	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
819	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
818	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
817	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
816	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
815	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
814	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
813	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
812	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
811	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
810	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
809	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
808	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
807	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
806	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
805	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
804	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
803	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
802	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
801	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
800	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
799	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
798	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
797	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
796	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
795	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
794	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
793	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
792	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
791	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
790	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
789	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
788	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
787	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
786	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
785	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
784	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
783	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
782	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
781	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
780	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
779	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
778	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
777	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
776	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
775	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
774	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
773	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
772	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
771	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
770	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
769	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
768	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
767	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
766	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
765	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
764	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
763	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
762	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
761	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
760	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
759	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
758	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
757	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
756	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
755	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
754	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
753	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
752	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
751	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
750	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
749	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
748	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
747	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
746	5	0	5	562	R15366	Human HSP70 protein.	4.21e+007
745	5	0	5	562	R15366	Human HSP70 protein.	4.21

[illegible]


```

DB WP137-731066/32.
N-PSD0: 732689.
PT New genes involved in sugar biosynthesis and attachment - used to
PT glycosylate antibodies, antitumorals etc. with altered pattern of
PT glycosylation.
PT Dislocation: Fig 4a: 58pp; English.
CC EryAII (M9736) is an enzyme involved in the biosynthesis of the
CC sugar L-lyxose. It is one of 10 enzymes (see also M9734-35 and
CC M9737-38) that are involved in the biosynthesis of the sugar
CC L-lyxose and attachment that have been identified from gene
CC clusters (732688 and 732685) of Saccharopolyspora erythraea.
CC Glycosylation-modified polynucleotides are produced by selectively
CC encoding a heterologous system supplementing the eryA and eryC gene
CC products with genes encoding glycosyltransferases and introducing them into
CC polynucleotide-producing microorganisms.
SQ Sequence 333 AA.
Query Match 0.7% Score 7; DB 23; Length 333;
Best Local Similarity 100.0%; Pred. No. 1 60+;+1;
Matches: 7; Conservative 0; Mismatches 0; Indels 0; Gaps:
DB 15 111111 21
Cy 637 ALTZERL 643
SEQUENCE 5
ID R45744 standard; Protein: 334 AA.
AC R45744:
DC 05-AUG-1994 (first entry)
DT Molsotolol dehydrogenase; production; engineering;
NM E.coli; substrate specificity; Bacillus.
OS Bacillus sp.
PI J06007158-A.
PI 23-JUN-1992: 16458.
PI 23-JUN-1992: JP-184546.
PA (ASNA) ASANI CHEN IND CO LTD.
DB WP137-4 033551/07.
PT Novel E.coli transformant containing molsotolol dehydrogenase
PT gene. for production and engineering of molsotolol
PT dehydrogenase.
PT Claim 3 page 8,9. jbp. Japanese.
CC molsotolol dehydrogenase plasmid which encodes a
CC molsotolol dehydrogenase. The plasmid can be transformed into a
CC suitable host thus producing molsotolol dehydrogenase quickly and
CC easily. 334 AA.
SQ Sequence

```

Mon Dec 21 09:28:22 1998

US-00-951-733-14.2a

Page 27

NM	Cytokine response protein; CR3; Interleukin-7; IL-7;
NP	Lymphocyte-stimulated gene expression; diagnosis; therapy;
OS	Homo sapiens.
PI	M09639427-AL.
PP	12 DEC -1986.
PR	05-JUN-1995: US-4611379.
PR	05-JUN-1995: US-4655883.
PR	05-JUN-1995: US-4623137.
PR	05-JUN-1995: US-4621082.
PR	05-JUN-1995: US-4630764.
PA	(DART-) DARKMOUTH COLLEGE.
PI	Breadling C.; Smith M.A.
DB	N-PIDB: T41372
PT	Cytokine response proteins and genes - used in the detection and
PT	therapy of diseases caused by a mutation in the CR coding region
PC	Claim 4, page 11-18; RHPH English; (WO81131-10) are encoded by
CC	genes (see also T41372-8) isolated from a cDNA library.
CC	Interleukin-7-induced human T-cell blast cDNA library. 6 Genes
CC	(CRI. 2, 3, 5, 6, 8) are novel. CR3 above homology to C-compiled
CC	and C-compiled sequences. The CR3 gene encodes a protein which
CC	probably plays a regulatory role in cellular proliferation and/or
CC	inflammation. It provides a novel receptor that allows the
CC	manipulation of cellular functions controlled by biochemical
CC	pathways signalled by the receptor and may be used to identify
CC	antibodies or drugs which modulate the function of the receptor
CC	Inflammation. Recombinant CR3 polypeptides can be produced in host
CC	cells e.g. as fusion proteins.
SO	Sequence 356 MA.
QW	Query Match
BE	Best Local Similarity 100.0%; Prod. No. 3, 804-01.
MT	Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
DB	245 PGMARC 251
YQ	241 PGMARC 247
II	
AC	W26169 standard; Protein: 375 AA.
AC	W26169_1987 (ELECT. PROT.)
DE	Human CD3 epsilon linked recitifier K channel-3 (XIRK-3).
KM	XIRK-3: kidney inward rectifier potassium ion channel; human.
OS	Homo sapiens.
PI	M09115912-A1.

```

Query Match          0.74 Score 7: DB 9: Length 334:
Beat Local Similarity 100.0% Pred. No. 3:80e+01:
Matches          7: Conservative      0: Mismatches      0: Indels      0: Gaps

DB      98 attlegu 104
          |||||
Query      304 KTSLEDA 310

RESULT      6
ID      M4246 standard: Protein: 358 AA.
CD      M3447: 1998 (first entry)
DB      Human NP4 prostaglandin receptor.
RM      Numm: NP4 prostaglandin receptor; adenylyate cyclase; drug screening
KM      CAMP.
PM      U5571835-1.
PF      10-725-1998.
PI      05-NAU-1994; 219431.
PI      05-NAU-1994; 219431.
PI      (ALTER.) ALTERNA INC.
PI      GLL DM, Regen DM, Woodward DF:
DR      WPI: 98-144807/3.
DR      DMR: 98-144807/3.
DR      DR: 98-144807/3.
CC      binding human NP4 prostaglandin receptor - useful for drug
PT      screening
PI      Claim 3: Column 17-20: 15pp: English.
CC      The present sequence represents human NP4 prostaglandin receptor.
CC      Expression vector, can be used to screen for substances that
CC      binding to the NP4 receptor; for substances that inhibit ligand
CC      binding to the NP4 receptor; and for NP4 receptor agonists
CC      transduction; increased production in cells pretreated with a
CC      Sequence 358 AA.: inhibition).
SQ      Sequence 358 AA.:

Query Match          0.74 Score 7: DB 20: Length 358:
Beat Local Similarity 100.0% Pred. No. 3:80e+01:
Matches          7: Conservative      0: Mismatches      0: Indels      0: Gaps

DB      245 p9arr1g 251
          |||||
Query      241 PCARNO 247

RESULT      7
ID      M08135 standard: Protein: 358 AA.
CD      M08135:
DB      11-NAU-1997 (first entry)
DR      Human cytokine response protein C93.

```

MON DEC 21 09:26:22 1999

US-08-951-733-14.Fa

Page 21

PD 10-APR-1987; UN4134.
PP 16-SEP-1985; US-003839.
PA (PNUA.) PHARMACIA & UPJOHN CO.
P1 BLENKOWSKI NJ /20.
PI 10-DEC-1987; UN4134.
DM 10-DEC-1987; UN4134.
PT New DNA encoding new human kidney potassium ion channel - and translocated cells, useful for screening for potassium channel modulators.

CC Page 31-32: 169P: English.
CC This polypeptide sequence comprises a human ATP-dependent potassium channel designated Kirr-3 (kidney inward rectifier K channel-3) that is expressed in the kidney, pancreas and lung but not brain.
CC The Kirr-3 sequence was deduced from cDNA clones (cKirr3 cDNA as probe, human Kirr-2 (see W6166) was also identified). Also
CC claimed are vector or plasmids containing Kirr-2 or Kirr-3 DNA,
CC and bacterially, mammalian or yeast cells containing such plasmids.
CC modifies human kidney, ATPase or related, K channel activity.
SQ Sequence 315 A%:

Query Match 0.7%; Score 7; DB 34; Length 315;
Matched Local Similarity 100.0%; Pvec No. J380+01; Caps
Matches 0; Mismatches 0; Indels 0;

Dn 260 repeated 266
Oy 804 TRIPAD 810

RESULZ 0
ID R08088 standard; Protein: 395 AA.
AC R08088:
DT 21-DIC-1995 (first entry)
DE galk gene of S. typhimurium gal operon
OS Streptococcus lividiana, serotype: gal⁻; gal⁻ operon: galactokinase.
PN US8435720-A.
PR 23-JUN-1995; 843706
PR 28-FEB-1986; US-0541706
PR 30-JAN-1987; US-005415
PR 29-APR-1991; US-693769
PR 19-OCT-1993; US-095979
PR 19-OCT-1993; US-095979
PA (ABVW) BLAWER M E.
PA (PONW) POERNALD J A.
PA (EDWD) SCHLITZ RJ.

DB 160 wvwhas 166
QV 767 AYVQMA 773

RESULT 14
ID R87508 standard: Protein: 493 AA.
AC R87508: (first entry)
DT 15-OCT-1995: (first entry)
DE Jojoba, *Simmondsia chinensis*,
NM Jojoba, fatty acyl-reductase; enzyme: transgenic plant; wax ester:
OS *Simmondsia chinensis*.
PM 09-DEC-1994: US-762556.
PR 01-JUN-1994: US-251464.
PA (CALJ) CALGENE INC.
DR N-PDB: 706749/3.
PT - used in the production of wax ester in plant cells not naturally
PS Produced using a plasmid p: 78pp. English.
CC This wild-type fatty-acyl-reductase may be reengineered such
CC 511) (706750). The sequence may be expressed in *Brassica* sp. used
CC to produce new oil and wax compositions containing a 44:2 wax ester
CC as a predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

RESULT 15
ID R26898 standard: Protein: 493 AA.
AC R26898: (first entry)
DT 15-FEB-1993: (first entry)
DE Jojoba, fatty acyl-reductase.
NM Jojoba, fatty acyl-reductase.
OS *Simmondsia chinensis*.
PM 03-SEP-1993.

CC new oil and wax compositions containing a 44:2 wax ester as a
CC predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

RESULT 17
ID R7163 standard: Protein: 493 AA.
AC R7163: (first entry)
DT 04-DEC-1995: (first entry)
DE Jojoba, fatty acyl-reductase.
NM Very long chain fatty acid: acyl-CoA: fatty acyl-reductase;
OS *Simmondsia chinensis*.
PM 08-JUN-1995.
PR 30-NOV-1994: US-766093.
PR 20-NOV-1991: US-762556.
PA (CALJ) CALGENE INC.
PI Lardilabai RD; Lesner MW, Metz JG;
DR N-PDB: 706749/3.
PT - used in the production of wax ester in plant cells not naturally
PS Produced using a plasmid p: 78pp. English.
CC This wild-type fatty-acyl-reductase may be reengineered such
CC 511) (706750). The sequence may be expressed in *Brassica* sp. used
CC to produce new oil and wax compositions containing a 44:2 wax ester
CC as a predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

PR 21-FEB-1993: US-762556.
PR 27-SEP-1991: US-762556.
PR 20-NOV-1991: US-762556.
PI Metz JG; Polard NR;
DR N-PDB: 706749/3.
PT - used in the production of wax ester in plant cells not naturally
PS Produced using a plasmid p: 78pp. English.
CC This wild-type fatty-acyl-reductase may be reengineered such
CC 511) (706750). The sequence may be expressed in *Brassica* sp. used
CC to produce new oil and wax compositions containing a 44:2 wax ester
CC as a predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

RESULT 16
ID R87509 standard: Protein: 493 AA.
AC R87509: (first entry)
DT 15-OCT-1995: (first entry)
DE Jojoba, fatty acyl-reductase.
NM Jojoba, fatty acyl-reductase; enzyme: transgenic plant; wax ester:
OS *Simmondsia chinensis*.
PM 09-DEC-1994: US-762556.
PR 01-JUN-1994: US-251464.
PA (CALJ) CALGENE INC.
DR N-PDB: 706749/3.
PT - used in the production of wax ester in plant cells not naturally
PS Produced using a plasmid p: 78pp. English.
CC This wild-type fatty-acyl-reductase may be reengineered such
CC 511) (706750). The sequence may be expressed in *Brassica* sp. used
CC to produce new oil and wax compositions containing a 44:2 wax ester
CC as a predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

ID R79934 standard: Protein: 493 AA.
AC R79934: (first entry)
DT 16-JAN-1996: (first entry)
DE Jojoba, fatty acyl-reductase.
NM Jojoba, fatty acyl-reductase; enzyme: transgenic plant; wax ester:
OS *Simmondsia chinensis*.
PM 05-DEC-1995: US-762556.
PR 20-NOV-1991: US-762556.
PR 21-NOV-1992: US-913411.
PR 13-NOV-1993: US-762556.
PR 20-NOV-1991: US-762556.
PA (CALJ) CALGENE INC.
PI Lardilabai RD; Lesner MW, Metz JG;
DR N-PDB: 706749/3.
PT - used in the production of wax ester in plant cells not naturally
PS Produced using a plasmid p: 78pp. English.
CC This wild-type fatty-acyl-reductase may be reengineered such
CC 511) (706750). The sequence may be expressed in *Brassica* sp. used
CC to produce new oil and wax compositions containing a 44:2 wax ester
CC as a predominant component.
SQ Sequence 493 AA:

Query Match
Best Local Similarity 100.0%; Pred. No. 3,804+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 426 111r1v4 432
QV 884 LLLRLVD 890

RESULT 18
ID R76470 standard: Protein: 493 AA.
AC R76470: (first entry)
DT 14-NOV-1995: (first entry)
DE Jojoba, fatty acyl-CoA reductase.
NM Jojoba, fatty acyl-CoA reductase; enzyme: transgenic plant; wax ester:
OS *Simmondsia chinensis*.
PM 05-DEC-1994: US-762556.
PR 21-FEB-1993: US-762556.
PR 27-SEP-1991: US-762556.
PR 20-NOV-1991: US-762556.

CC is used in the transactivating factor inhibitor of the invention, that
CC inhibit Tax-dependent transcription. The inhibitors can be used to
CC inhibit transcription of DNA that codes for proteins that can be harmful
CC to humans, especially human, e.g. to treat viral infections (e.g.
CC hepatitis B, hepatitis C, HIV, AIDS, cytomegalovirus (CMV),
CC human immunodeficiency virus (HIV), cytomegalovirus (CMV),
CC sequence 15 AA:

Query Match 0.64: Score 6; DB 28; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.71e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 profile 10
Qy 365 PEPIDS 170

RESULT 34
ID W42350 standard: Peptide: 15 AA.

AC W42350 (first entry)
DE 08-APR-1998 (first entry)
NM Bacteriophage peptide library: peptide epitope: therapeutic target:
OS Synthetic compound library: Interleukin-8; 15-0.
PS MO9735194-A1.
PI 25-SEP-1997; 004136.
PA (HARD) HARVARD COLLEGE.
PI 97-480355/44.
DR Identifying compounds which interact with target molecules - using
PT selected compounds, target molecules and testing of enantiomers of
PS Disclosure: Fig 6: 89pp: English.
CC 15-mer peptides W42350-60 are part of a peptide library. They were
CC identified as potential inhibitors of the interaction between a target
CC which interact with a target molecule, and comprises contacting a
CC screening molecule comprising solid target molecule, or the enantiomer of
CC with the target molecule are selected, and the ability of their
CC enantiomer to interact with the target molecule is tested. Ligands for a
CC target protein can be identified by combining a D-enantiomer of a target
CC protein (a D-target protein), and a variegated compound library, and
CC desired binding interaction with the D-target protein. The methods can
CC be used for identifying agonists or antagonists of targets such as
CC receptors, enzymes, DNA binding proteins or signal transduction

CC approach in addition to scoring for interaction of functional groups.
CC They provide a powerful selection method that allows for the production
CC of ligands with the same diversity as peptides but with the greatly
CC improved pharmacokinetic profiles needed for drug activity.
SQ Sequence 15 AA:

Query Match 0.64: Score 6; DB 27; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.71e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 affral 7
Qy 67 MAFRAL 72

RESULT 36
ID R61310 standard: peptide: 15 AA.

AC R61310 (first entry)
DE 25-APR-1995 (first entry)
NM HIV transactivating protein: Tax; regulator: host cell: transcription;
OS protein-protein interaction: inhibition: virus: regulation:
PS Human T-cell leukemia virus.
PI 97-424972/39.
PI 97-424972/39.
PA (HARD) HARVARD COLLEGE.
PI 97-424972/39.
DR Identifying compounds which interact with target molecules - using
PT selected compounds, target molecules and testing of enantiomers of
PS Disclosure: Fig 6: 89pp: English.
CC 15-mer peptides W42350-60 are part of a peptide library. They were
CC identified as potential inhibitors of the interaction between a target
CC which interact with a target molecule, and comprises contacting a
CC screening molecule comprising solid target molecule, or the enantiomer of
CC with the target molecule are selected, and the ability of their
CC enantiomer to interact with the target molecule is tested. Ligands for a
CC target protein can be identified by combining a D-enantiomer of a target
CC protein (a D-target protein), and a variegated compound library, and
CC desired binding interaction with the D-target protein. The methods can
CC be used for identifying agonists or antagonists of targets such as
CC receptors, enzymes, DNA binding proteins or signal transduction

Query Match 0.64: Score 6; DB 11; Length 15;

CC proteins. The methods can provide a structurally selective
CC approach in addition to scoring for interaction of functional groups.
CC They provide a powerful selection method that allows for the production
CC of ligands with the same diversity as peptides but with the greatly
CC improved pharmacokinetic profiles needed for drug activity.
SQ Sequence 15 AA:

Query Match 0.64: Score 6; DB 27; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.71e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2 affral 7
Qy 67 MAFRAL 72

RESULT 35
ID W42350 standard: Peptide: 15 AA.

AC W42350 (first entry)
DE 08-APR-1998 (first entry)
NM Bacteriophage peptide library: peptide epitope: therapeutic target:
OS Synthetic compound library: Interleukin-8; 15-0.
PS MO9735194-A1.
PI 25-SEP-1997; 004136.
PA (HARD) HARVARD COLLEGE.
PI 97-480355/44.
DR Identifying compounds which interact with target molecules - using
PT selected compounds, target molecules and testing of enantiomers of
PS Disclosure: Fig 6: 89pp: English.
CC 15-mer peptides W42350-60 are part of a peptide library. They were
CC identified as potential inhibitors of the interaction between a target
CC which interact with a target molecule, and comprises contacting a
CC screening molecule comprising solid target molecule, or the enantiomer of
CC with the target molecule are selected, and the ability of their
CC enantiomer to interact with the target molecule is tested. Ligands for a
CC target protein can be identified by combining a D-enantiomer of a target
CC protein (a D-target protein), and a variegated compound library, and
CC desired binding interaction with the D-target protein. The methods can
CC be used for identifying agonists or antagonists of targets such as
CC receptors, enzymes, DNA binding proteins or signal transduction

CC approach in addition to scoring for interaction of functional groups.
CC They provide a powerful selection method that allows for the production
CC of ligands with the same diversity as peptides but with the greatly
CC improved pharmacokinetic profiles needed for drug activity.
SQ Sequence 15 AA:

Query Match 0.64: Score 6; DB 27; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.71e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 profile 10
Qy 365 PEPIDS 170

RESULT 37
ID W42350 standard: peptide: 16 AA.

AC W42350 (first entry)
DE 27-MAR-1998 (first entry)
NM Bacteriophage peptide library: peptide epitope: therapeutic target:
OS Synthetic compound library: Interleukin-8; 15-0.
PS MO9735194-A1.
PI 25-SEP-1997; 004136.
PA (HARD) HARVARD COLLEGE.
PI 97-480355/44.
DR Identifying compounds which interact with target molecules - using
PT selected compounds, target molecules and testing of enantiomers of
PS Disclosure: Fig 6: 89pp: English.
CC 16-mer peptides W42350-61 are part of a peptide library. They were
CC identified as potential inhibitors of the interaction between a target
CC which interact with a target molecule, and comprises contacting a
CC screening molecule comprising solid target molecule, or the enantiomer of
CC with the target molecule are selected, and the ability of their
CC enantiomer to interact with the target molecule is tested. Ligands for a
CC target protein can be identified by combining a D-enantiomer of a target
CC protein (a D-target protein), and a variegated compound library, and
CC desired binding interaction with the D-target protein. The methods can
CC be used for identifying agonists or antagonists of targets such as
CC receptors, enzymes, DNA binding proteins or signal transduction

Query Match 0.64: Score 6; DB 26; Length 16;

CC cultured to produce the protein. The alpha-glucuronidase protein is used
CC to improve dough and baked products. The protein can also be used to
CC treat cellulose pulp to remove at least some of the glucuronic acid
CC residue present in it, as a slugging additive and to degrade plant
CC matter generally. When added to animal feed particularly the protein
CC increases feed utilization and feed conversion rate, particularly by
CC increasing bioavailability of metal ions.

50 Sequence 35 AA:

Query Match 0.68: Score 6; DB 28; Length 35;

Best Local Similarity 85.78; Pred. No. 4, 71e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 5 align 11

07 857 S1STLL 863

RESULT 50

ID W20450 standard: Protein: 45 AA.

AC W20450: 11-JUN-1997 (first entry)

DT 11-JUN-1997 (first entry)

RT 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

RV 11-JUN-1997 (first entry)

CC acid sequences predicted from various ORF were analyzed for significant
CC homology to other known or exported membrane proteins. Having identified
CC and determined the sequences of interest, particular regions can be
CC isolated from R. pylori for application for recombinant polypeptide
CC production in R. coli hosts.

50 Sequence 45 AA:

Query Match 0.68: Score 6; DB 22; Length 45;

Best Local Similarity 100.00; Pred. No. 7, 71e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 8 align 13

07 682 M58VU 687

Search completed: Fri Dec 18 18:56:12 1998

Job time : 210 secs.

WU

Release 3.1A John F. Collins, Bioinformatics Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

protein - protein database search, using Smith-Waterman algorithm
Run on: Fri Dec 18 18:47:50 1998; Maspar time 11.24 seconds
1037.201 Million cell updates/sec

Tabular output not generated.

Title: >US-08-951-733-14

Description: (1.949) from US08951733.rpx

Perfect Score: 949

Sequence: 1 BASCONCYLUTRANALPMT.....PRDNDLCCGACVONPANC 949

Scoring table: TABLE uninterpretable

Gap 60

Searched: 107076 seqs, 3441958 residues

Post-processing: Minimum Match 0%

Database: 1:plrt 2:plrt 3:plrt 4:plrt

Statistics: Mean 3.996; Variance 0.453; scale 8.838

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description	Prod. No.
1	9	0.9	24	2	crinamembrane glycopr	9.74e-04

1	9	0.9	183	3	A49315	leukocyte common Ant1	8.74e-04
2	9	0.9	183	3	A49315	leukocyte common Ant1	9.74e-04
3	8	0.8	169	2	A35521	protein kinase (EC 2.7.1.37)	1.29e-01
4	8	0.8	305	1	D48559	ureacil-PN6 glycosylas	1.29e-01
5	8	0.8	342	1	TYH495	protein kinase (EC 2.7.1.37)	1.29e-01
6	8	0.8	400	1	ORF42R	protein kinase (EC 2.7.1.37)	1.29e-01
7	8	0.8	400	1	ORF42R	protein kinase (EC 2.7.1.37)	1.29e-01
8	8	0.8	508	2	JC6300	cyclohexane P450oxc -	1.29e-01
9	8	0.8	654	1	HMR551	dnaf-type molecular c	1.29e-01
10	8	0.8	970	2	S39069	inositol polyphosphat	1.29e-01
11	8	0.9	1480	2	S48410	poly(h)-specific fibro	1.02e-01
12	8	0.9	39	2	CO0082	hypothetical 4.3K pro	1.04e-01
13	8	0.7	53	2	SO1945	myosin catalytic ligh	1.04e-01
14	7	0.7	53	2	SO1945	myosin catalytic ligh	1.04e-01
15	7	0.7	53	2	SO1945	myosin catalytic ligh	1.04e-01
16	7	0.7	53	2	SO1945	myosin catalytic ligh	1.04e-01
17	7	0.7	53	2	SO1945	myosin catalytic ligh	1.04e-01
18	7	0.7	111	2	S41936	hypothetical protein	1.04e-01
19	7	0.7	111	2	S41936	hypothetical protein	1.04e-01
20	7	0.7	128	2	S77068	hypothetical protein	1.04e-01
21	7	0.7	139	1	RHW07A	T-cell receptor alpha	1.04e-01
22	7	0.7	148	2	S10380	hper protein - fasher	1.04e-01
23	7	0.7	168	2	S64196	hypothetical protein	1.04e-01
24	7	0.7	177	2	JC5478	endoxerythronuclease	1.04e-01
25	7	0.7	182	2	S14240	yo20 protein - yeast	1.04e-01
26	7	0.7	203	1	MOR33V	myosin alkali light c	1.04e-01
27	7	0.7	203	1	MOR33V	myosin alkali light c	1.04e-01
28	7	0.7	209	1	G69784	hypothetical protein	1.04e-01
29	7	0.7	267	2	S76610	hypothetical protein	1.04e-01
30	7	0.7	278	2	G01327	alpha-cooperetol tran	1.04e-01
31	7	0.7	278	2	S54352	alpha-cooperetol tran	1.04e-01
32	7	0.7	280	2	B59137	conserved hypothetical	1.04e-01
33	7	0.7	307	2	S75008	dnaf protein - synch	1.04e-01
34	7	0.7	316	2	S16681	homeotic protein - nu	1.04e-01
35	7	0.7	321	2	H64877	peptide transport sys	1.04e-01
36	7	0.7	321	2	H64877	peptide transport sys	1.04e-01
37	7	0.7	321	2	S62316	egg protein phosphon	1.04e-01
38	7	0.7	321	2	S62316	egg protein phosphon	1.04e-01
39	7	0.7	329	2	S58158	hypothetical protein	1.04e-01
40	7	0.7	342	2	PC4111	hepatocellular carcin	1.04e-01
41	7	0.7	342	2	S56795	hypothetical protein	1.04e-01
42	7	0.7	348	2	S19450	ubiquinone 4-epimeras	1.04e-01
43	7	0.7	358	2	S51112	EP2 prostaflandin rec	1.04e-01
44	7	0.7	388	2	S53087	ubiquinol--cyclohone	1.04e-01
45	7	0.7	395	1	K15MG	galactokinase (EC 2.7.1.6)	1.04e-01
46	7	0.7	423	2	S61142	metacytic form spec1	1.04e-01
47	7	0.7	423	2	S61142	metacytic form spec1	1.04e-01
48	7	0.7	430	2	S53125	endo-beta-1,6-glucan	1.04e-01
49	7	0.7	448	2	A59076	nitrogenase molybden	1.04e-01
50	7	0.7	448	2	A59076	nitrogenase molybden	1.04e-01
51	7	0.7	465	2	S69755	hypothetical protein	1.04e-01
52	7	0.7	465	2	S69755	hypothetical protein	1.04e-01

Mon Dec 21 09:28:23 1998

US-08-951-793-14.rpt

Page

155	6	0.6	145	2	533556	proteasome membrane protein	33e02
156	6	0.6	147	2	Ar756	snmp core protein d1	33e02
157	6	0.6	148	2	Ar756	snmp core protein d1	33e02
158	6	0.6	149	2	J65054	riboses-5-phosphate 3-epimerase	33e02
159	6	0.6	149	2	J65054	riboses-5-phosphate 3-epimerase	33e02
160	6	0.6	150	2	535516	ecmectolase (EC 3.1.1)	33e02
161	6	0.6	152	2	535516	hypothetical protein	33e02
162	6	0.6	153	2	535516	hypothetical protein	33e02
163	6	0.6	154	2	Ar6967	transposase homolog	33e02
164	6	0.6	153	1	GGGCR	glutathion - horn shell (-	33e02
165	6	0.6	160	2	535913	Allylgren RALF precursor	33e02
166	6	0.6	161	2	CG1853	Id4 Nuclein	33e02
167	6	0.6	161	2	CG1853	Id4 Nuclein	33e02
168	6	0.6	161	2	533660	helix-loop-helix prot	33e02
169	6	0.6	163	2	53606	hypothetical protein	33e02
170	6	0.6	163	2	53606	hypothetical protein	33e02
171	6	0.6	166	1	NP02	vasopressin / neurop	33e02
172	6	0.6	166	1	NP02	vasopressin / neurop	33e02
173	6	0.6	166	1	NP02	vasopressin / neurop	33e02
174	6	0.6	166	1	NP02	vasopressin / neurop	33e02
175	6	0.6	166	1	NP02	vasopressin / neurop	33e02
176	6	0.6	166	1	NP02	vasopressin / neurop	33e02
177	6	0.6	167	1	533442	lutealising hormone b	33e02
178	6	0.6	167	1	533442	lutealising hormone b	33e02
179	6	0.6	167	1	533442	lutealising hormone b	33e02
180	6	0.6	168	1	NP02	vasopressin / neurop	33e02
181	6	0.6	168	1	NP02	vasopressin / neurop	33e02
182	6	0.6	172	2	G70334	conserved hypothetical fibroblast growth fac	33e02
183	6	0.6	172	2	G70334	conserved hypothetical fibroblast growth fac	33e02
184	6	0.6	172	2	G70334	conserved hypothetical fibroblast growth fac	33e02
185	6	0.6	172	2	G70334	conserved hypothetical fibroblast growth fac	33e02
186	6	0.6	172	2	G70334	conserved hypothetical fibroblast growth fac	33e02
187	6	0.6	176	2	532704	Leptin17-C230 prote	33e02
188	6	0.6	176	2	532704	Leptin17-C230 prote	33e02
189	6	0.6	176	2	532704	Leptin17-C230 prote	33e02
190	6	0.6	177	2	B49054	cytochrome c oxidase subunit1/AdhA my1	33e02
191	6	0.6	177	2	B49054	cytochrome c oxidase subunit1/AdhA my1	33e02
192	6	0.6	181	2	536432	T-cell receptor beta	33e02
193	6	0.6	181	2	536432	T-cell receptor beta	33e02
194	6	0.6	184	2	53205	hypothetical protein	33e02
195	6	0.6	184	2	53205	hypothetical protein	33e02
196	6	0.6	185	2	53205	hypothetical protein	33e02
197	6	0.6	185	2	53205	hypothetical protein	33e02
198	6	0.6	185	2	53205	hypothetical protein	33e02
199	6	0.6	189	2	536930	adenylylate cyclase hoc	33e02
200	6	0.6	189	2	536930	adenylylate cyclase hoc	33e02
201	6	0.6	191	2	536930	adenylylate cyclase hoc	33e02
202	6	0.6	191	2	536930	adenylylate cyclase hoc	33e02
203	6	0.6	191	2	536930	adenylylate cyclase hoc	33e02
204	6	0.6	191	2	536930	adenylylate cyclase hoc	33e02
205	6	0.6	199	2	535521	probable ribosomal pr	33e02
206	6	0.6	199	2	535521	probable ribosomal pr	33e02
207	6	0.6	199	2	535521	probable ribosomal pr	33e02
208	6	0.6	199	2	535521	probable ribosomal pr	33e02
209	6	0.6	199	2	535521	probable ribosomal pr	33e02
210	6	0.6	199	2	535521	probable ribosomal pr	33e02
211	6	0.6	199	2	535521	probable ribosomal pr	33e02

[illegible]

257	6	0	5	244	1	A10686	hematophyl juvenile ho	33e+02
258	6	0	5	245	1	S14422	hypothetical protein	33e+02
259	6	0	5	246	1	S14423	hypothetical protein	33e+02
260	6	0	5	247	2	D59004	hypothetical protein	33e+02
261	6	0	5	248	2	J05032	Katexaurin-B - rTiclos	33e+02
262	6	0	5	249	2	G54463	cobalamins (γ'-phospho	33e+02
263	6	0	5	250	2	G54464	cobalamins (γ'-phospho	33e+02
264	6	0	5	251	2	G54465	cobalamins (γ'-phospho	33e+02
265	6	0	5	252	2	J30358	hypothetical protein	33e+02
266	6	0	5	253	2	J30358	HR-73 protein - sea a	33e+02
267	6	0	5	254	2	C59172	conserved hypothetical	33e+02
268	6	0	5	255	2	G69109	conserved hypothetical	33e+02
269	6	0	5	256	1	R1871	Thyrosinase precursor	33e+02
270	6	0	5	257	2	S13135	alcohol dehydrogenase	33e+02
271	6	0	5	257	2	S13135	thymidine deaminase	33e+02
272	6	0	5	258	2	A11131	phosphotransfer center	33e+02
273	6	0	5	259	2	A11131	phosphotransfer center	33e+02
274	6	0	5	260	2	S12956	hypothetical protein	33e+02
275	6	0	5	260	2	J02367	UDP-N-acetylglucosamyl	33e+02
276	6	0	5	261	2	S19355	entry exclusion prote	33e+02
277	6	0	5	262	2	B6445	ThyB - class II histidyl	33e+02
278	6	0	5	263	2	B6445	ThyB - class II histidyl	33e+02
279	6	0	5	264	2	B6445	Fleischlar absent adhe	33e+02
280	6	0	5	265	2	B70137	phosphatase ABC transpo	33e+02
281	6	0	5	266	2	B70137	transcription initiatin	33e+02
282	6	0	5	267	2	B64131	hypothetical protein	33e+02
283	6	0	5	268	2	J05540	2-hydroxypropan-3-ate	33e+02
284	6	0	5	269	2	A13738	Endopl protein - sea	33e+02
285	6	0	5	269	2	A13738	hypothetical 29.4 kD	33e+02
286	6	0	5	270	2	F64139	NR skin precursor (C	33e+02
287	6	0	5	271	2	S12753	hypothetical protein	33e+02
288	6	0	5	272	2	A15661	tyrosinase (EC 1.4.2.15)	33e+02
289	6	0	5	273	2	J04171	tyrosinase (EC 1.4.2.15)	33e+02
290	6	0	5	274	2	A12110	tyrosinase (EC 1.4.2.15)	33e+02
291	6	0	5	275	2	A12110	tyrosinase (EC 1.4.2.15)	33e+02
292	6	0	5	276	2	G43459	surv survival protein	33e+02
293	6	0	5	276	2	G43459	surv survival protein	33e+02
294	6	0	5	277	2	S09662	hypothetical protein	33e+02
295	6	0	5	278	2	S19371	hypothetical protein	33e+02
296	6	0	5	279	2	S19371	hypothetical protein	33e+02
297	6	0	5	280	2	S16384	hypothetical protein	33e+02
298	6	0	5	281	2	D39660	Balmain ring 6 chain	33e+02
299	6	0	5	282	2	J07062	pre-mRNA splicing fac	33e+02
300	6	0	5	283	2	S16161	hypothetical protein	33e+02
301	6	0	5	284	2	A18536	hypothetical protein	33e+02
302	6	0	5	285	2	G44479	sol atropine / chro	33e+02
303	6	0	5	286	2	B13270	E2 membrane glycoprot	33e+02
304	6	0	5	287	2	S17447	mark protein pax6o	33e+02
305	6	0	5	288	2	S17447	mark protein pax6o	33e+02
306	6	0	5	289	2	S14951	hypothetical protein	33e+02
307	6	0	5	290	2	S21498	fertilin precursor (C	33e+02)

108	6	0.6	286	3	608939	signal sequence, recep	4.3e-007
109	6	0.6	286	3	500010	analogous to 3'-acyl	4.3e-007
110	6	0.6	286	3	500010	analogous to 3'-acyl	4.3e-007
111	6	0.6	286	3	500010	analogous to 3'-acyl	4.3e-007
112	6	0.6	286	3	500010	analogous to 3'-acyl	4.3e-007
113	6	0.6	286	3	500010	analogous to 3'-acyl	4.3e-007
114	6	0.6	289	1	MMV017	11.7K protein - tobac	4.3e-007
115	6	0.6	289	1	MMV017	11.7K protein - tobac	4.3e-007
116	6	0.6	289	1	MMV017	11.7K protein - tobac	4.3e-007
117	6	0.6	289	1	MMV017	11.7K protein - tobac	4.3e-007
118	6	0.6	292	2	AS119	equipoirin 3 - human	4.3e-007
119	6	0.6	292	2	AS119	equipoirin 3 - human	4.3e-007
120	6	0.6	292	2	AS119	equipoirin 3 - human	4.3e-007
121	6	0.6	292	2	AS119	equipoirin 3 - human	4.3e-007
122	6	0.6	300	1	500922	hypothetical protein	4.3e-007
123	6	0.6	300	1	500922	hypothetical protein	4.3e-007
124	6	0.6	300	1	500922	hypothetical protein	4.3e-007
125	6	0.6	303	3	579933	hypothetical protein	4.3e-007
126	6	0.6	303	3	579933	hypothetical protein	4.3e-007
127	6	0.6	304	2	513892	modulation protein no	4.3e-007
128	6	0.6	304	2	513892	modulation protein no	4.3e-007
129	6	0.6	305	1	OM5054	onlinase catalyze	4.3e-007
130	6	0.6	305	2	700030	conserved hypothetical	4.3e-007
131	6	0.6	306	2	506559	band 3 anion transpor	4.3e-007
132	6	0.6	306	2	506559	band 3 anion transpor	4.3e-007
133	6	0.6	305	2	506559	hypothetical protein	4.3e-007
134	6	0.6	307	2	504080	carrier protein, WCI,	4.3e-007
135	6	0.6	310	2	B51121	Chloroplast Import as	4.3e-007
136	6	0.6	310	2	150696	unknown - chicken (fr	4.3e-007
137	6	0.6	311	2	522550	hypothetical protein	4.3e-007
138	6	0.6	311	2	522550	hypothetical protein	4.3e-007
139	6	0.6	312	1	DCG437	uvelin-DNA glycosyl	4.3e-007
140	6	0.6	312	2	CG1487	glucan endo-1,3-beta-	4.3e-007
141	6	0.6	312	2	CG1487	glucan endo-1,3-beta-	4.3e-007
142	6	0.6	312	2	CG1487	glucan endo-1,3-beta-	4.3e-007
143	6	0.6	312	2	CG1487	glucan endo-1,3-beta-	4.3e-007
144	6	0.6	314	1	PNM010	beta-lactamase (EC 3	4.3e-007
145	6	0.6	314	1	PNM010	beta-lactamase (EC 3	4.3e-007
146	6	0.6	315	2	549694	cysteine synthase, 1S	4.3e-007
147	6	0.6	316	2	523943	cysteine synthase, 1S	4.3e-007
148	6	0.6	317	2	567573	ferrioxalase, NAC trans	4.3e-007
149	6	0.6	317	2	567573	M032 homolog A05.ort	4.3e-007
150	6	0.6	317	2	549493	Lab02 protein - strep	4.3e-007
151	6	0.6	318	2	CG1653	malloproteinase, 1 (f	4.3e-007
152	6	0.6	319	2	528897	myogenic factor 3 - h	4.3e-007
153	6	0.6	319	2	528897	myogenic factor 3 - h	4.3e-007
154	6	0.6	322	2	A68081	NMDA dehydrogenase, a	4.3e-007
155	6	0.6	322	2	A68081	NMDA dehydrogenase, a	4.3e-007
156	6	0.6	324	2	544400	hypothetical protein	4.3e-007
157	6	0.6	324	2	544400	hypothetical protein	4.3e-007
158	6	0.6	326	1	579956	hypothetical protein	4.3e-007
159	6	0.6	326	1	PRWV04	proteases (EC 3.4.2)	4.3e-007

553 6 0 6 483 2 B69780 transretinoyl retin
554 6 0 6 483 2 119719 nitrogen regulatory p
555 6 0 6 486 1 MYSL10 aszrayan (EC 3.4.24
556 6 0 6 487 1 B65310 probable membrane pro
557 6 0 6 489 1 DRYD20 dopamine receptor D1,
558 6 0 6 490 2 844170 C-14 sterol reductase
559 6 0 6 490 2 cytochrome P450 2C18
560 6 0 6 491 2 A51272 pre-B cell enhancing
561 6 0 6 492 2 J05149 alkaline nucleases (EC
562 6 0 6 493 2 A32454 X-Pro dipeptidase (EC
563 6 0 6 493 2 A32454 probable general secr
564 6 0 6 494 2 B64324 hypothetical protein
565 6 0 6 497 2 C02838 cytochrome P450 - era
566 6 0 6 497 2 04H0D1 debrisoquine 4-hydrox
567 6 0 6 502 2 A51597 Mstork-Aldrich syndr
568 6 0 6 503 2 488174 Mstork-Aldrich syndr
569 6 0 6 504 2 S10451 nuclear protein, 25k
570 6 0 6 504 2 S17248 regulatory protein M1
571 6 0 6 504 2 A37215 g111, cellie masting
572 6 0 6 505 2 S33962 endoglucanase - Ewin
573 6 0 6 506 2 C64377 hypothetical protein
574 6 0 6 506 2 POLG31 gag polypeptide - sam
575 6 0 6 509 2 669581 alkyl hydroperoxide r
576 6 0 6 511 2 A46638 cholesteryl monooxyge
577 6 0 6 514 2 S21377 unspecific monooxygen
578 6 0 6 519 2 669888 hydroxymethylglutaryl
579 6 0 6 521 1 A43135 chromaffin granule am
580 6 0 6 521 2 S55677 glutamate transporter
581 6 0 6 523 2 S27612 ketoglutarate semiald
582 6 0 6 525 2 S03600 cell surface antigen
583 6 0 6 527 1 S84022 cell surface antigen
584 6 0 6 528 2 A41718 polyuridylic tract-
585 6 0 6 529 2 B24059 t complex polypeptide
586 6 0 6 529 2 A47131 Na+-dependent neutral
587 6 0 6 530 2 B23151 sermidin (ary) ABC
588 6 0 6 530 2 S15553 polyuridylic tract-
589 6 0 6 530 2 521015 gene 40 protein - equ
590 6 0 6 531 2 S21015 polyuridylic tract-

591 6 0 6 593 2 S65470 privative deacetylase
592 6 0 6 600 2 A33658 pirinallier rod pro
593 6 0 6 602 2 669182 cyclooxigenase C/H syn
594 6 0 6 602 2 639782 cyclooxigenase 1 - ra
595 6 0 6 603 2 S28544 computational factor XI
596 6 0 6 604 2 S28544 computational factor XI
597 6 0 6 604 2 669793 glutamine-fructose-6
598 6 0 6 604 2 S25203 amn protein - Strept
599 6 0 6 606 2 A40313 maternal effect prote
600 6 0 6 607 2 S25432 conserved hypothetical
601 6 0 6 607 2 S25432 conserved hypothetical
602 6 0 6 608 1 ABOH52 serum albumin 2 precu
603 6 0 6 610 2 641115 hypothetical protein
604 6 0 6 612 2 S53114 hypothetical protein
605 6 0 6 615 2 P64128 proteinase IV homolog
606 6 0 6 616 1 PRVGLL probable RNA-directed
607 6 0 6 617 2 A51646 5-methyltetrahydrofol
608 6 0 6 619 2 A40646 transforming protein
609 6 0 6 620 2 S56790 probable membrane pro
610 6 0 6 621 2 S35092 plakatolobin - mouse
611 6 0 6 625 2 S35871 Pael 17 protein - mou
612 6 0 6 629 2 A59814 ABC transporter (ATP-
613 6 0 6 630 2 S37111 DNA topoisomerase IV
614 6 0 6 631 2 152257 epistatin - mouse
615 6 0 6 632 2 B42603 ADP-beta-glucoside-
616 6 0 6 632 2 S41572 DNA topoisomerase IV
617 6 0 6 632 2 S41572 DNA topoisomerase IV
618 6 0 6 632 2 S41572 DNA topoisomerase IV
619 6 0 6 632 2 S41572 DNA topoisomerase IV
620 6 0 6 632 2 S41572 DNA topoisomerase IV
621 6 0 6 632 2 S41572 DNA topoisomerase IV
622 6 0 6 632 2 S41572 DNA topoisomerase IV
623 6 0 6 632 2 S41572 DNA topoisomerase IV
624 6 0 6 632 2 S41572 DNA topoisomerase IV
625 6 0 6 632 2 S41572 DNA topoisomerase IV
626 6 0 6 632 2 S41572 DNA topoisomerase IV
627 6 0 6 632 2 S41572 DNA topoisomerase IV
628 6 0 6 632 2 S41572 DNA topoisomerase IV
629 6 0 6 632 2 S41572 DNA topoisomerase IV
630 6 0 6 632 2 S41572 DNA topoisomerase IV
631 6 0 6 632 2 S41572 DNA topoisomerase IV
632 6 0 6 632 2 S41572 DNA topoisomerase IV
633 6 0 6 632 2 S41572 DNA topoisomerase IV
634 6 0 6 632 2 S41572 DNA topoisomerase IV
635 6 0 6 632 2 S41572 DNA topoisomerase IV
636 6 0 6 632 2 S41572 DNA topoisomerase IV
637 6 0 6 632 2 S41572 DNA topoisomerase IV
638 6 0 6 632 2 S41572 DNA topoisomerase IV
639 6 0 6 632 2 S41572 DNA topoisomerase IV
640 6 0 6 632 2 S41572 DNA topoisomerase IV
641 6 0 6 632 2 S41572 DNA topoisomerase IV
642 6 0 6 632 2 S41572 DNA topoisomerase IV
643 6 0 6 632 2 S41572 DNA topoisomerase IV
644 6 0 6 632 2 S41572 DNA topoisomerase IV
645 6 0 6 632 2 S41572 DNA topoisomerase IV
646 6 0 6 632 2 S41572 DNA topoisomerase IV
647 6 0 6 632 2 S41572 DNA topoisomerase IV
648 6 0 6 632 2 S41572 DNA topoisomerase IV
649 6 0 6 632 2 S41572 DNA topoisomerase IV
650 6 0 6 632 2 S41572 DNA topoisomerase IV
651 6 0 6 632 2 S41572 DNA topoisomerase IV
652 6 0 6 632 2 S41572 DNA topoisomerase IV
653 6 0 6 632 2 S41572 DNA topoisomerase IV
654 6 0 6 632 2 S41572 DNA topoisomerase IV
655 6 0 6 632 2 S41572 DNA topoisomerase IV
656 6 0 6 632 2 S41572 DNA topoisomerase IV
657 6 0 6 632 2 S41572 DNA topoisomerase IV
658 6 0 6 632 2 S41572 DNA topoisomerase IV
659 6 0 6 632 2 S41572 DNA topoisomerase IV
660 6 0 6 632 2 S41572 DNA topoisomerase IV
661 6 0 6 632 2 S41572 DNA topoisomerase IV
662 6 0 6 632 2 S41572 DNA topoisomerase IV
663 6 0 6 632 2 S41572 DNA topoisomerase IV
664 6 0 6 632 2 S41572 DNA topoisomerase IV
665 6 0 6 632 2 S41572 DNA topoisomerase IV
666 6 0 6 632 2 S41572 DNA topoisomerase IV
667 6 0 6 632 2 S41572 DNA topoisomerase IV
668 6 0 6 632 2 S41572 DNA topoisomerase IV
669 6 0 6 632 2 S41572 DNA topoisomerase IV
670 6 0 6 632 2 S41572 DNA topoisomerase IV
671 6 0 6 632 2 S41572 DNA topoisomerase IV
672 6 0 6 632 2 S41572 DNA topoisomerase IV
673 6 0 6 632 2 S41572 DNA topoisomerase IV
674 6 0 6 632 2 S41572 DNA topoisomerase IV
675 6 0 6 632 2 S41572 DNA topoisomerase IV
676 6 0 6 632 2 S41572 DNA topoisomerase IV
677 6 0 6 632 2 S41572 DNA topoisomerase IV
678 6 0 6 632 2 S41572 DNA topoisomerase IV
679 6 0 6 632 2 S41572 DNA topoisomerase IV
680 6 0 6 632 2 S41572 DNA topoisomerase IV
681 6 0 6 632 2 S41572 DNA topoisomerase IV
682 6 0 6 632 2 S41572 DNA topoisomerase IV
683 6 0 6 632 2 S41572 DNA topoisomerase IV
684 6 0 6 632 2 S41572 DNA topoisomerase IV
685 6 0 6 632 2 S41572 DNA topoisomerase IV
686 6 0 6 632 2 S41572 DNA topoisomerase IV
687 6 0 6 632 2 S41572 DNA topoisomerase IV
688 6 0 6 632 2 S41572 DNA topoisomerase IV
689 6 0 6 632 2 S41572 DNA topoisomerase IV
690 6 0 6 632 2 S41572 DNA topoisomerase IV
691 6 0 6 632 2 S41572 DNA topoisomerase IV
692 6 0 6 632 2 S41572 DNA topoisomerase IV
693 6 0 6 632 2 S41572 DNA topoisomerase IV
694 6 0 6 632 2 S41572 DNA topoisomerase IV
695 6 0 6 632 2 S41572 DNA topoisomerase IV
696 6 0 6 632 2 S41572 DNA topoisomerase IV
697 6 0 6 632 2 S41572 DNA topoisomerase IV
698 6 0 6 632 2 S41572 DNA topoisomerase IV
699 6 0 6 632 2 S41572 DNA topoisomerase IV
700 6 0 6 632 2 S41572 DNA topoisomerase IV
701 6 0 6 632 2 S41572 DNA topoisomerase IV
702 6 0 6 632 2 S41572 DNA topoisomerase IV
703 6 0 6 632 2 S41572 DNA topoisomerase IV
704 6 0 6 632 2 S41572 DNA topoisomerase IV
705 6 0 6 632 2 S41572 DNA topoisomerase IV
706 6 0 6 632 2 S41572 DNA topoisomerase IV
707 6 0 6 632 2 S41572 DNA topoisomerase IV
708 6 0 6 632 2 S41572 DNA topoisomerase IV
709 6 0 6 632 2 S41572 DNA topoisomerase IV
710 6 0 6 632 2 S41572 DNA topoisomerase IV
711 6 0 6 632 2 S41572 DNA topoisomerase IV
712 6 0 6 632 2 S41572 DNA topoisomerase IV
713 6 0 6 632 2 S41572 DNA topoisomerase IV
714 6 0 6 632 2 S41572 DNA topoisomerase IV
715 6 0 6 632 2 S41572 DNA topoisomerase IV

614 6 0 6 511 2 J70511 muscarinic acetylcholin
615 6 0 6 512 2 B69123 2-oxoacid ferredoxin
616 6 0 6 512 2 S56652 probable calcium-depe
617 6 0 6 513 2 S44058 TPA1 protein-related
618 6 0 6 513 2 S44058 TPA1 protein-related
619 6 0 6 517 2 B46535 Interleukin-9 recepto
620 6 0 6 518 1 F0MNM gag polypeptide - Mol
621 6 0 6 519 2 A35052 Interleukin-2 recept
622 6 0 6 519 2 A35052 Interleukin-2 recept
623 6 0 6 519 2 A35052 Interleukin-2 recept
624 6 0 6 519 2 A35052 Interleukin-2 recept
625 6 0 6 519 2 A35052 Interleukin-2 recept
626 6 0 6 519 2 A35052 Interleukin-2 recept
627 6 0 6 519 2 A35052 Interleukin-2 recept
628 6 0 6 519 2 A35052 Interleukin-2 recept
629 6 0 6 519 2 A35052 Interleukin-2 recept
630 6 0 6 519 2 A35052 Interleukin-2 recept
631 6 0 6 519 2 A35052 Interleukin-2 recept
632 6 0 6 519 2 A35052 Interleukin-2 recept
633 6 0 6 519 2 A35052 Interleukin-2 recept
634 6 0 6 519 2 A35052 Interleukin-2 recept
635 6 0 6 519 2 A35052 Interleukin-2 recept
636 6 0 6 519 2 A35052 Interleukin-2 recept
637 6 0 6 519 2 A35052 Interleukin-2 recept
638 6 0 6 519 2 A35052 Interleukin-2 recept
639 6 0 6 519 2 A35052 Interleukin-2 recept
640 6 0 6 519 2 A35052 Interleukin-2 recept
641 6 0 6 519 2 A35052 Interleukin-2 recept
642 6 0 6 519 2 A35052 Interleukin-2 recept
643 6 0 6 519 2 A35052 Interleukin-2 recept
644 6 0 6 519 2 A35052 Interleukin-2 recept
645 6 0 6 519 2 A35052 Interleukin-2 recept
646 6 0 6 519 2 A35052 Interleukin-2 recept
647 6 0 6 519 2 A35052 Interleukin-2 recept
648 6 0 6 519 2 A35052 Interleukin-2 recept
649 6 0 6 519 2 A35052 Interleukin-2 recept
650 6 0 6 519 2 A35052 Interleukin-2 recept
651 6 0 6 519 2 A35052 Interleukin-2 recept
652 6 0 6 519 2 A35052 Interleukin-2 recept
653 6 0 6 519 2 A35052 Interleukin-2 recept
654 6 0 6 519 2 A35052 Interleukin-2 recept
655 6 0 6 519 2 A35052 Interleukin-2 recept
656 6 0 6 519 2 A35052 Interleukin-2 recept
657 6 0 6 519 2 A35052 Interleukin-2 recept
658 6 0 6 519 2 A35052 Interleukin-2 recept
659 6 0 6 519 2 A35052 Interleukin-2 recept
660 6 0 6 519 2 A35052 Interleukin-2 recept
661 6 0 6 519 2 A35052 Interleukin-2 recept
662 6 0 6 519 2 A35052 Interleukin-2 recept
663 6 0 6 519 2 A35052 Interleukin-2 recept
664 6 0 6 519 2 A35052 Interleukin-2 recept
665 6 0 6 519 2 A35052 Interleukin-2 recept
666 6 0 6 519 2 A35052 Interleukin-2 recept
667 6 0 6 519 2 A35052 Interleukin-2 recept
668 6 0 6 519 2 A35052 Interleukin-2 recept
669 6 0 6 519 2 A35052 Interleukin-2 recept
670 6 0 6 519 2 A35052 Interleukin-2 recept
671 6 0 6 519 2 A35052 Interleukin-2 recept
672 6 0 6 519 2 A35052 Interleukin-2 recept
673 6 0 6 519 2 A35052 Interleukin-2 recept
674 6 0 6 519 2 A35052 Interleukin-2 recept
675 6 0 6 519 2 A35052 Interleukin-2 recept
676 6 0 6 519 2 A35052 Interleukin-2 recept
677 6 0 6 519 2 A35052 Interleukin-2 recept
678 6 0 6 519 2 A35052 Interleukin-2 recept
679 6 0 6 519 2 A35052 Interleukin-2 recept
680 6 0 6 519 2 A35052 Interleukin-2 recept
681 6 0 6 519 2 A35052 Interleukin-2 recept
682 6 0 6 519 2 A35052 Interleukin-2 recept
683 6 0 6 519 2 A35052 Interleukin-2 recept
684 6 0 6 519 2 A35052 Interleukin-2 recept
685 6 0 6 519 2 A35052 Interleukin-2 recept
686 6 0 6 519 2 A35052 Interleukin-2 recept
687 6 0 6 519 2 A35052 Interleukin-2 recept
688 6 0 6 519 2 A35052 Interleukin-2 recept
689 6 0 6 519 2 A35052 Interleukin-2 recept
690 6 0 6 519 2 A35052 Interleukin-2 recept
691 6 0 6 519 2 A35052 Interleukin-2 recept
692 6 0 6 519 2 A35052 Interleukin-2 recept
693 6 0 6 519 2 A35052 Interleukin-2 recept
694 6 0 6 519 2 A35052 Interleukin-2 recept
695 6 0 6 519 2 A35052 Interleukin-2 recept
696 6 0 6 519 2 A35052 Interleukin-2 recept
697 6 0 6 519 2 A35052 Interleukin-2 recept
698 6 0 6 519 2 A35052 Interleukin-2 recept
699 6 0 6 519 2 A35052 Interleukin-2 recept
700 6 0 6 519 2 A35052 Interleukin-2 recept
701 6 0 6 519 2 A35052 Interleukin-2 recept
702 6 0 6 519 2 A35052 Interleukin-2 recept
703 6 0 6 519 2 A35052 Interleukin-2 recept
704 6 0 6 519 2 A35052 Interleukin-2 recept
705 6 0 6 519 2 A35052 Interleukin-2 recept
706 6 0 6 519 2 A35052 Interleukin-2 recept
707 6 0 6 519 2 A35052 Interleukin-2 recept
708 6 0 6 519 2 A35052 Interleukin-2 recept
709 6 0 6 519 2 A35052 Interleukin-2 recept
710 6 0 6 519 2 A35052 Interleukin-2 recept
711 6 0 6 519 2 A35052 Interleukin-2 recept
712 6 0 6 519 2 A35052 Interleukin-2 recept
713 6 0 6 519 2 A35052 Interleukin-2 recept
714 6 0 6 519 2 A35052 Interleukin-2 recept
715 6 0 6 519 2 A35052 Interleukin-2 recept

716 6 0 6 437 2 S51603 dent-type molecular c
717 6 0 6 437 2 B64202 type IIS restriction
718 6 0 6 683 1 A41785 system b(0+) amino a
719 6 0 6 687 2 A39045 protein-glutamine gam
720 6 0 6 690 2 A36143 phospholipase C (EC
721 6 0 6 692 2 A36143 phospholipase C (EC
722 6 0 6 694 2 S78444 di27 protein - fruit
723 6 0 6 695 2 S64300 amphiphysin (clone 22
724 6 0 6 695 2 S64300 amphiphysin (clone 22
725 6 0 6 697 2 B44489 CR binding protein
726 6 0 6 701 2 S64489 Tbx1 protein - mouse
727 6 0 6 701 2 S64489 Tbx1 protein - mouse
728 6 0 6 704 2 S64489 Tbx1 protein - mouse
729 6 0 6 704 2 S64489 Tbx1 protein - mouse
730 6 0 6 704 2 S64489 Tbx1 protein - mouse
731 6 0 6 704 2 S64489 Tbx1 protein - mouse
732 6 0 6 704 2 S64489 Tbx1 protein - mouse
733 6 0 6 704 2 S64489 Tbx1 protein - mouse
734 6 0 6 704 2 S64489 Tbx1 protein - mouse
735 6 0 6 704 2 S64489 Tbx1 protein - mouse
736 6 0 6 704 2 S64489 Tbx1 protein - mouse
737 6 0 6 704 2 S64489 Tbx1 protein - mouse
738 6 0 6 704 2 S64489 Tbx1 protein - mouse
739 6 0 6 704 2 S64489 Tbx1 protein - mouse
740 6 0 6 704 2 S64489 Tbx1 protein - mouse
741 6 0 6 704 2 S64489 Tbx1 protein - mouse
742 6 0 6 704 2 S64489 Tbx1 protein - mouse
743 6 0 6 704 2 S64489 Tbx1 protein - mouse
744 6 0 6 704 2 S64489 Tbx1 protein - mouse
745 6 0 6 704 2 S64489 Tbx1 protein - mouse
746 6 0 6 704 2 S64489 Tbx1 protein - mouse
747 6 0 6 704 2 S64489 Tbx1 protein - mouse
748 6 0 6 704 2 S64489 Tbx1 protein - mouse
749 6 0 6 704 2 S64489 Tbx1 protein - mouse
750 6 0 6 704 2 S64489 Tbx1 protein - mouse
751 6 0 6 704 2 S64489 Tbx1 protein - mouse
752 6 0 6 704 2 S64489 Tbx1 protein - mouse
753 6 0 6 704 2 S64489 Tbx1 protein - mouse
754 6 0 6 704 2 S64489 Tbx1 protein - mouse
755 6 0 6 704 2 S64489 Tbx1 protein - mouse
756 6 0 6 704 2 S64489 Tbx1 protein - mouse
757 6 0 6 704 2 S64489 Tbx1 protein - mouse
758 6 0 6 704 2 S64489 Tbx1 protein - mouse
759 6 0 6 704 2 S64489 Tbx1 protein - mouse
760 6 0 6 704 2 S64489 Tbx1 protein - mouse
761 6 0 6 704 2 S64489 Tbx1 protein - mouse
762 6 0 6 704 2 S64489 Tbx1 protein - mouse
763 6 0 6 704 2 S64489 Tbx1 protein - mouse
764 6 0 6 704 2 S64489 Tbx1 protein - mouse
765 6 0 6 704 2 S64489 Tbx1 protein - mouse
766 6 0 6 704 2 S64489 Tbx1 protein - mouse
767 6 0 6 704 2 S64489 Tbx1 protein - mouse
768 6 0 6 704 2 S64489 Tbx1 protein - mouse
769 6 0 6 704 2 S64489 Tbx1 protein - mouse
770 6 0 6 704 2 S64489 Tbx1 protein - mouse
771 6 0 6 704 2 S64489 Tbx1 protein - mouse
772 6 0 6 704 2 S64489 Tbx1 protein - mouse
773 6 0 6 704 2 S64489 Tbx1 protein - mouse
774 6 0 6 704 2 S64489 Tbx1 protein - mouse
775 6 0 6 704 2 S64489 Tbx1 protein - mouse
776 6 0 6 704 2 S64489 Tbx1 protein - mouse
777 6 0 6 704 2 S64489 Tbx1 protein - mouse
778 6 0 6 704 2 S64489 Tbx1 protein - mouse
779 6 0 6 704 2 S64489 Tbx1 protein - mouse
780 6 0 6 704 2 S64489 Tbx1 protein - mouse
781 6 0 6 704 2 S64489 Tbx1 protein - mouse
782 6 0 6 704 2 S64489 Tbx1 protein - mouse
783 6 0 6 704 2 S64489 Tbx1 protein - mouse
784 6 0 6 704 2 S64489 Tbx1 protein - mouse
785 6 0 6 704 2 S64489 Tbx1 protein - mouse
786 6 0 6 704 2 S64489 Tbx1 protein - mouse
787 6 0 6 704 2 S64489 Tbx1 protein - mouse
788 6 0 6 704 2 S64489 Tbx1 protein - mouse
789 6 0 6 704 2 S64489 Tbx1 protein - mouse
790 6 0 6 704 2 S64489 Tbx1 protein - mouse
791 6 0 6 704 2 S64489 Tbx1 protein - mouse
792 6 0 6 704 2 S64489 Tbx1 protein - mouse
793 6 0 6 704 2 S64489 Tbx1 protein - mouse
794 6 0 6 704 2 S64489 Tbx1 protein - mouse
795 6 0 6 704 2 S64489 Tbx1 protein - mouse
796 6 0 6 704 2 S64489 Tbx1 protein - mouse
797 6 0 6 704 2 S64489 Tbx1 protein - mouse
798 6 0 6 704 2 S64489 Tbx1 protein - mouse
799 6 0 6 704 2 S64489 Tbx1 protein - mouse
800 6 0 6 704 2 S64489 Tbx1 protein - mouse
801 6 0 6 704 2 S64489 Tbx1 protein - mouse
802 6 0 6 704 2 S64489 Tbx1 protein - mouse
803 6 0 6 704 2 S64489 Tbx1 protein - mouse
804 6 0 6 704 2 S64489 Tbx1 protein - mouse
805 6 0 6 704 2 S64489 Tbx1 protein - mouse
806 6 0 6 704 2 S64489 Tbx1 protein - mouse
807 6 0 6 704 2 S64489 Tbx1 protein - mouse
808 6 0 6 704 2 S64489 Tbx1 protein - mouse
809 6 0 6 704 2 S64489 Tbx1 protein - mouse
810 6 0 6 704 2 S64489 Tbx1 protein - mouse
811 6 0 6 704 2 S64489 Tbx1 protein - mouse
812 6 0 6 704 2 S64489 Tbx1 protein - mouse
813 6 0 6 704 2 S64489 Tbx1 protein - mouse
814 6 0 6 704 2 S64489 Tbx1 protein - mouse
815 6 0 6 704 2 S64489 Tbx1 protein - mouse
816 6 0 6 704 2 S64489 Tbx1 protein - mouse
817 6 0 6 704 2 S64489 Tbx1 protein - mouse
818 6 0 6 704 2 S64489 Tbx1 protein - mouse
819 6 0 6 704 2 S64489 Tbx1 protein - mouse
820 6 0 6 704 2 S64489 Tbx1 protein - mouse
821 6 0 6 704 2 S64489 Tbx1 protein - mouse
822 6 0 6 704 2 S64489 Tbx1 protein - mouse
823 6 0 6 704 2 S64489 Tbx1 protein - mouse
824 6 0 6 704 2 S64489 Tbx1 protein - mouse
825 6 0 6 704 2 S64489 Tbx1 protein - mouse
826 6 0 6 704 2 S64489 Tbx1 protein - mouse
827 6 0 6 704 2 S64489 Tbx1 protein - mouse
828 6 0 6 704 2 S64489 Tbx1 protein - mouse
829 6 0 6 704 2 S64489 Tbx1 protein - mouse
830 6 0 6 704 2 S64489 Tbx1 protein - mouse
831 6 0 6 704 2 S64489 Tbx1 protein - mouse
832 6 0 6 704 2 S64489 Tbx1 protein - mouse
833 6 0 6 704 2 S64489 Tbx1 protein - mouse
834 6 0 6 704 2 S64489 Tbx1 protein - mouse
835 6 0 6 704 2 S64489 Tbx1 protein - mouse
836 6 0 6 704 2 S64489 Tbx1 protein - mouse
837 6 0 6 704 2 S64489 Tbx1 protein - mouse
838 6 0 6 704 2 S64489 Tbx1 protein - mouse
839 6 0 6 704 2 S64489 Tbx1 protein - mouse
840 6 0 6 704 2 S64489 Tbx1 protein - mouse
841 6 0 6 704 2 S64489 Tbx1 protein - mouse
842 6 0 6 704 2 S64489 Tbx1 protein - mouse
843 6 0 6 704 2 S64489 Tbx1 protein - mouse
844 6 0 6 704 2 S64489 Tbx1 protein - mouse
845 6 0 6 704 2 S64489 Tbx1 protein - mouse
846 6 0 6 704 2 S64489 Tbx1 protein - mouse
847 6 0 6 704 2 S64489 Tbx1 protein - mouse
848 6 0 6 704 2 S64489 Tbx1 protein - mouse
849 6 0 6 704 2 S64489 Tbx1 protein - mouse
850 6 0 6 704 2 S64489 Tbx1 protein - mouse
851 6 0 6 704 2 S64489 Tbx1 protein - mouse

[illegible]

```

across-reference NCBI 9006605
accession NCID R100682
molecule-type protein
residues 78-342 label POT

REFERENCE
PUBMED 1A15740
AUTHOR Taylor A.R., Taylor S.S.
J Biol Chem (1980) 255(68):8488
TITLE Covalent modification of an adenosine 3',5'-monophosphate binding site of the regulatory subunit of cAMP-dependent protein kinase II with 8-azidoadenosine
PMID 61800682

across-reference NCID R100682
accession A15740
molecule-type protein
label NCBI 9006605
COMMENT This catalytic form of the enzyme is composed of two regulatory chains and two catalytic chains. Activation by cAMP produces four active catalytic monomers and a regulatory dimer that binds two cAMP molecules.

COMMENT From cAMP-modulatory chains are found: Tsalpha, Ibeta, II-alpha, and II-beta. Their expression varies among tissues and is in some cases constitutive and in others inducible.

COMMENT Type II regulatory chain are phosphorylated after activated phosphotransferase by other kinases as unclear.

COMMENT Type II regulatory chains mediate membrane association by binding to anchoring proteins, including the MAP Kinase. The phosphorylation of serine residues is essential for their binding and for dimerization.

CLASSIFICATION superfamily cAMP-dependent protein kinase regulatory chain; cAMP receptor protein cyclic nucleotide-binding domain
KEYWORDS cAMP binding duplication; heterotrimer; homodimer; phosphoprotein; phosphotransferase

FEATURES
REGION 81-135
domain cAMP receptor protein cyclic nucleotide-binding
SUMMARY
length 169 aminoacids 135
Query Match 0.84; Score 8; DB: 2; Length 165;
Database Match Similarity 10.0; Posid 128.0;
Machine Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 52 RECDTOR 59
QY 461 RECDTOR 468
RESULT 5

```

[illegible]

```

CMTGCTCGTCAAGTCAGACATGGTTTGCCAGCATGCCTGGGAAAGGCGAAGAAGAGGTGGGTA
Agene      mos
CLASSIFICATION    specifically Kinase-related transforming protein; protein kinase homology; phosphotransferase; serpin/rheonine-specific
KEYWORDS          protein kinase; transforming protein
FEATURE           61-347            adenin protein kinase homology label KIN
                  68-77             region protein kinase Appending motif
SIMONARY          length 312 ;molecular-weight 37070 ;sequence 8888
Query Match      0 %; Score 8; Dn 1; Length 342;
Meet Local Similarity Maximal 0; Conservative 0; Identical 0; Gaps 0;
Matches 81 Conservative
Db       291 LabeledLGA 298
Cy       J12 UniProtKB   J19
RESULT        7
TITLE         OREGOR     type complete
AUTHOR        regulatory kinases (EC 2.7.1.1); cAMP-dependent, cyclic II-alpha
ORGANISM      regulatory chain - b-type
DATE          15-Oct-1992
REFERENCE     A00618:D:617058
AUTHOR        A00618
            Makiio R.; Smith G.B.; Reaba E.C.; Walsh K.A.; Titani, K. et al.
            Biochemistry 31(4) 1420-1406
ABSTRACT      Amino acid sequence of the regulatory subunit of bovine tyrosyl
            II adenosine cyclase 3',5'-phosphate dependent protein kinase.
cross-reference NC00619
accession     A00618
molculcile_type protein
**residues 617058 61708 : Valine; TAK
journal       Arch. Biochem. Biophys. (1991) 289:187-191
title        Phosphorylation of RIi subunit and attenuation of
            cAMP-dependent protein Kinase activity by proline-directed
            kinase
cross-reference CMBD:6178531
accession     #617058
molculcile_type protein
**residues 15-166 :Alahab; BNA
COMMENT      The inactive form of the enzyme is composed of two regulatory

```

Choline and two catalytic chains. Activation by cAMP produces two active catalytic monomers and a regulatory dimer that binds four cAMP molecules.

COMMENT Four types of regulatory chains are found: I-alpha, I-beta, II-alpha, and II-beta. Their expression varies among tissues and is regulated by the cAMP levels. Type II regulatory chains are phosphorylated by the activated catalytic chain. The physiological significance of phosphorylation by other kinases is unclear.

COMMENT Type II regulatory chains mediate membrane association by binding amino-terminal 50 residues are essential for this binding and for dimerization.

CLASSIFICATION Superfamily cAMP-dependent protein kinase regulatory chain: homology amino end; cAMP binding; duplication; heterotrimer; phosphoprotein.

KEYWORDS

FEATURE 1-114
115-266
237-369
48-211
95
204-213
314-343

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 106 EEXDTPR 113
QY 461 EEXDTPR 468

RESULT 8
ENTRY ORK028
TITLE protein kinase (EC 2.7.1.37), cAMP-dependent, type II-alpha
ORGANISM *Escherichia coli* (strain ATCC 8739) common name man
DATE 07-Sep-1997

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 110 EEXDTPR 117
QY 461 EEXDTPR 468

RESULT 9
ENTRY JG6300
TITLE cytochrome P-450c1 - chicken
ORGANISM *Gallus gallus gallus* common name chicken
DATE 11-Dec-1997

ACCESSIONS JG6300
REFERENCE Kurebayashi, O.; Mihimori, K.; Matsuo, S.
TITLE The cDNA cloning and transient expression of a chicken gene encoding cytochrome P-450c1.
FEATURES
FEATURE 1-508
FEATURE 509-588
FEATURE 589-650

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 128 PRCVPLMT 135
QY 426 PRCVPLMT 433

RESULT 10
ENTRY HHR51
TITLE DNA type molecular map of a precursor, mitochondrial
ORGANISM *Escherichia coli* (strain ATCC 8739) common name man
DATE 07-Sep-1997

ACCESSIONS 503885
REFERENCE 503885
AUTHORS Oeyen, O.; Wylsbach, F.; Scott, J.D.; Hanson, V.; Jahnson, T.
TITLE RBS Letc. (1989) 246:37-44
COMMENT The inactive form of the enzyme is composed of two regulatory chains and two catalytic chains. Activation by cAMP produces two active catalytic monomers and a regulatory dimer that binds four cAMP molecules.

COMMENT Four types of regulatory chains are found: I-alpha, I-beta, II-alpha, and II-beta. Their expression varies among tissues and is regulated by the cAMP levels. Type II regulatory chains are phosphorylated by the activated catalytic chain. The physiological significance of phosphorylation by other kinases is unclear.

COMMENT Type II regulatory chains mediate membrane association by binding amino-terminal 50 residues are essential for this binding and for dimerization.

CLASSIFICATION Superfamily cAMP-dependent protein kinase regulatory chain: homology amino end; cAMP binding; duplication; heterotrimer; phosphoprotein.

KEYWORDS

FEATURE 2-404
2-118
139-260
261-393

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 106 EEXDTPR 113
QY 461 EEXDTPR 468

RESULT 8
ENTRY ORK028
TITLE protein kinase (EC 2.7.1.37), cAMP-dependent, type II-alpha
ORGANISM *Escherichia coli* (strain ATCC 8739) common name man
DATE 07-Sep-1997

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 110 EEXDTPR 117
QY 461 EEXDTPR 468

RESULT 9
ENTRY JG6300
TITLE cytochrome P-450c1 - chicken
ORGANISM *Gallus gallus gallus* common name chicken
DATE 11-Dec-1997

ACCESSIONS JG6300
REFERENCE Kurebayashi, O.; Mihimori, K.; Matsuo, S.
TITLE The cDNA cloning and transient expression of a chicken gene encoding cytochrome P-450c1.
FEATURES
FEATURE 1-508
FEATURE 509-588
FEATURE 589-650

SUMMARY
Query Match
Best Local Similarity 100.0%; Pred. No. 1,296-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 128 PRCVPLMT 135
QY 426 PRCVPLMT 433

RESULT 10
ENTRY HHR51
TITLE DNA type molecular map of a precursor, mitochondrial
ORGANISM *Escherichia coli* (strain ATCC 8739) common name man
DATE 07-Sep-1997

[illegible][illegible]

```

**sequences**
**accession** I-1480 **label** VOS
**references** EMBL:X79743
**authors**
Voss R., Zemann, A.; Teodoru, C.; Valencia, A.; Senese, C.;
Wiemers, S.; Schwyger, C.; Zimmermann, J.; Sender, C.;
Aebi, M., 11-61-78
**journal** Nucleotide sequence and analysis of the centromeric region o
**title** Yeast chromosome IX.
**accension**
**sequence** nucleic acid sequence not shown
**accession**
**sequence_type** 306-355;640-688 **label** VOW
**references** EMBL:X79743
**genetics**
**gene** GCD-PAN1, NMD3, NRP3
**map_position** 98
**map_coordinates** GCD:3000145; NRP3:170006
**keywords** Hydrolysis
**synonym** Hydrolyst 1480 amolecular-weight 160367 scheckman 2111
**molecular_weight** 160367
**length** 1480
**cd** 0 88; Procd 8; DB 2; Length 1480;
Best Local Similarity 100.0%; Spec.No. 1.29e+01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Caps 0;
|||||||
DB 1232 AVTPNAC 1239
|||||||
CY 439 AVTPNAC 446

RESULT 13
ENTRY
TITLE
AUTHOR
ORGANISM
DATE
ACCESSIONS
REFERENCE
JOURNAL
TITLE
cross-reference NMD3:92109718
accession JM0515
sequence_type protein
molecule_type protein
organismal source eukaryote
REFERENCE
AD7374
AUTHORS
Mol., Endocrinol. (1991) 5:938-948

```

[illegible]

ALTERNATE_NAMES myosin alkali light chain
 ORIGINISM formal_name Mus musculus scconmon_name mouse
 DATE 01-Dec-1998 sequence_revision 01-Dec-1998 text_change
 ACCSSIONS S01345
 REFERENCE Cohen, A.; Barton, P. J. R.; Robert, B.; Garner, I.; Alonso,
 #journal Nucleic Acids Res. 16(1988) 16:10037-10052
 #title Promoter analysis of myosin alkali light chain genes
 #comment expressed in mouse striated muscle.
 #cross-references MID:89057447
 #molecule-type DNA
 #residues 1-52 #label COH
 #accession EMBL:X12972
 #classification superfamily calmodulin repeat homology
 #keywords repeats
 #summary length 52 checksum 8566

Query Match 0.7% Score 7; DB 2; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 34 AMPEPR 40
 QY 264 AMPEPR 270

RESULT 16
 ENTRY A8196 #type fragment
 #title myosin alkali light chain p15 (fragment)
 #formal_name Cavla Porcellus scconmon_name guinea pig
 #ORIGINISM 19-Dec-1993 sequence_revision 18-Nov-1994 text_change
 DATE 06-Jun-1997
 ACCSSIONS A49136
 REFERENCE Roselli-Rehman, L.; Robbins, L. S.; Cone, R. D.
 #journal Endocrinology (1993) 130:1837-1861
 #title Thyrotropin receptor messenger ribonucleic acid is expressed
 #cross-references MID:9218789
 #accession A49136
 #status preliminary; not compared with conceptual translation
 #molecule-type nucleic acid
 #residues 1-105
 #comment experimental source adipose tissue
 #note sequence extracted from MCB1 backbone (NCBI:89464)
 #classification superfamily glycoprotein hormone receptor; leucine-rich
 alpha-2-glycoprotein repeat homology

across-references MID:8604365
 #accession A23329
 #molecule-type RNA
 #residues 1-115 #label SHE
 #cross-references MID:919819; MID:919820
 #classification superfamily immunoglobulin; leukocyte common
 antigen cytolellic domain homology;
 protein-tyrosine-phosphatase homology
 #keywords glycoprotein
 #summary length 115 checksum 1887

Query Match 0.7% Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 121 POFALLD 127
 QY 121 POFALLD 127

RESULT 19
 ENTRY S77068 #type complete
 #title hypothetical protein - synchocystis sp. (PCC 6803)
 #ORIGINISM PCC 6803
 #date 25-Apr-1997 sequence_revision 25-Apr-1997 text_change
 DATE 09-Sep-1997
 ACCSSIONS S77068
 REFERENCE Kondo, T.; Sato, S.; Kozaki, H.; Tanaka, A.; Asanuma, E.;
 Nakamura, Y.; Miyajima, M.; Hirose, M.; Sugita, M.;
 Sasamoto, S.; Kikuchi, T.; Hosouchi, T.; Matsumoto, A.;
 Muraki, A.; Nakaseki, M.; Haruo, K.; Okumura, S.; Shimo,
 Y.; Yada, K.; Tsubota, T.; Watanabe, A.; Yamada, M.
 #journal DNA Res. (1996) 3:109-116
 #title Sequence analysis of the genome of the unicellular
 cyanobacterium synchocystis sp. PCC6803. II. Sequence
 analysis of the synchocystis sp. PCC6803. II. Sequence
 potential protein-coding regions.
 #cross-references MID:97061201
 #accession S77068
 #status preliminary
 #molecule-type protein
 #residues 1-128 #label KAN
 #cross-references EMBL:D64005; MID:91001779; PID:91006607
 #note The nucleotide sequence was submitted to the EMBL Data
 #summary length 128 checksum 2173

Query Match 0.7% Score 7; DB 2; Length 128;

SYNARY length 52 checksum 4896
 Query Match 0.7% Score 7; DB 2; Length 52;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 46 LELVAN 52
 QY 99 LELVAN 105

RESULT 17
 ENTRY S14855 #type complete
 #title hypothetical protein 2 yeast (Kansensula polymorpha)
 #ORIGINISM 20-Feb-1995 sequence_revision 20-Feb-1995 text_change
 DATE 09-Sep-1997
 ACCSSIONS S14855
 REFERENCE Kuznetsov, A. I.; Seregin, S. A.; Tikhonova, L. P.; Krut'kov,
 #journal V.M. Kuznetsov, A. I.; Seregin, S. A.; Tikhonova, L. P.; Krut'kov,
 #submitted to the EMBL Data Library, April 1991
 #description Nucleotide sequence of Kansensula polymorpha DNA region
 #comment containing DNA.
 #accession S14855
 #status preliminary
 #molecule-type DNA
 #residues 1-111 #label KNU
 #cross-references MID:927762; PID:82764
 #summary length 111 checksum 3577

Query Match 0.7% Score 7; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 22 TSLAT 28
 QY 640 TSLAT 646

RESULT 18
 ENTRY A23329 #type fragment
 #title myosin alkali light chain p15 (fragment)
 #formal_name Mus musculus scconmon_name mouse
 #ORIGINISM 05-Jun-1987 sequence_revision 05-Jun-1987 text_change
 DATE 20-Mar-1998
 ACCSSIONS A23329
 REFERENCE Shen, F. W.; Suga, T.; Litman, G.; Freeman, G.; Tung, J. S.;
 Cantor, H.; Boyse, E. A.
 #journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7360-7363

Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 46 AYZIGS 52
 QY 811 AYZIGS 817

RESULT 20
 ENTRY A92150 #type complete
 #title T-cell receptor alpha chain precursor V region (CTV-1L7) -
 #ORIGINISM human
 #formal_name Homo sapiens scconmon_name man
 #date 17-Mar-1987 sequence_revision 17-Mar-1987 text_change
 DATE 09-Sep-1997
 ACCSSIONS A92150
 REFERENCE Laiden, J. M.; Fraser, J. D.; Strominger, J. L.
 #journal Immunogenetics (1986) 24:17-23
 #title The organization and structure of the T-cell receptor genes
 #comment from an alternative cytolellic human T-lymphocyte clone.
 #cross-references MID:8676770
 #accession A92150
 #molecule-type RNA
 #residues 1-129 #label LFI
 #comment This sequence was derived from a human cytotoxic T-lymphocyte that
 is T3+ T4+ T8-.

GENETICS
 #cross-references MID:910404; ONIX:186880
 #map position 14q11.2-14q11.3
 #classification superfamily immunoglobulin V region; immunoglobulin homology
 #keywords glycoprotein; heterotrimer; receptor; T-cell
 #summary length 129 checksum 1887

Query Match 0.7% Score 7; DB 1; Length 129;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 4 LIGASV 10

***residues 1-348 (label) RMS
 GENETICS
 #cross-reference EMBL:823476; NID:6396530; PID:6396531
 #gene exons
 CLASSIFICATION
 #superfamily UDPglucose 4-epimerase homology
 #EC 5.1.1.35
 #domain UDPglucose 4-epimerase homology label UDP
 #length 348 #molecular-weight 36700 #checksum 5393
 SUMMARY
 Query Match 0.7% Score 7; DB 2; Length 348;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 342 GCGASRS 348
 QY 247 GCGASRS 253

RESULT 43
 #type complete
 #protein prostaiglandin E2 receptor - human
 #formal_name Homo sapiens (common_name) man
 #date 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change
 #accessions 138920
 #references 138920
 #authors Regan, J.W.; Bailey, T.J.; Peppert, D.J.; Pierce, K.L.;
 Bogardus, A.M.; Donillo, J.E.; Falldorf, C.E.; Kozlowski,
 M.O. Pharmacol. (1994) 48:213-220
 #journal Cloning of a novel human prostaglandin receptor with
 characteristics of the pharmacologically defined EP2
 #title
 #cross-reference NID:634359403
 #accession 138920
 #status preliminary; translated from CG/EMBL/DBJ
 #molecule_type mRNA
 #residues 358 #label RMS
 #cross-reference EMBL:019407; NID:639719; PID:632650
 #length 358 #molecular-weight 36675 #checksum 150
 SUMMARY
 Query Match 0.7% Score 7; DB 2; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 245 PGNRRG 351
 QY 241 PGNRRG 247

b6 homology; plastocyanin reductase 1IK
 protein homology; electron transfer; heme; iron; mitochondrion;
 oxidative phosphorylation; oxidoreductases; respiratory
 chain; transmembrane protein
 FEATURE
 1-10 #domain cytochrome b homology label CBN
 10-210 #domain cytochrome b6 homology label CBN
 15-51 #domain transmembrane status predicted label TM1
 80-98 #domain transmembrane status predicted label TM2
 118-134 #domain transmembrane status predicted label TM3
 135-171 #domain transmembrane status predicted label TM4
 222-340 #domain plastocyanin-plastocyanin reductase 1IK protein
 homology label 1IK
 230-246 #domain transmembrane status predicted label TM5
 248-305 #domain transmembrane status predicted label TM6
 306-358 #domain transmembrane status predicted label TM7
 355-371 #domain transmembrane status predicted label TM8
 96-197 binding site heme iron, high potential (HIS) (axial
 ligands) status predicted
 #length 385 #molecular-weight 43192 #checksum 3308
 SUMMARY
 Query Match 0.7% Score 7; DB 2; Length 385;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 8 LLLALVD 14
 QY 884 LLLALVD 890

RESULT 46
 #type complete
 #protein galactokinase (EC 2.7.1.6) - Streptomyces sp.
 #formal_name Streptomyces sp. (common_name) man
 #date 28-Feb-1997 #sequence_revision 11-Dec-1990 #text_change
 #accessions C28669
 #references C28669
 #authors J. Bacteriol. (1988) 170:203-212
 #journal Gene organization and structure of the Streptomyces lividans
 #title
 #cross-reference EMBL:8808669
 #accession C28669
 #molecule_type DNA
 #residues 1-395 #label ADM
 #genetics the source is designated as Streptomyces lividans
 #gene galk

RESULT 44
 #type complete
 #protein EP2 prostaglandin receptor - human
 #formal_name Homo sapiens (common_name) man
 #date 10-Sep-1997 #sequence_revision 01-Sep-1995 #text_change
 #accessions 651112
 #references 651112
 #authors J. Biol. Chem. (1997) 272:1000-1005
 #journal Submitted to the EMBL Data Library, January 1995
 #title preliminary
 #status
 #molecule_type mRNA #label ORF
 #cross-reference EMBL:X83868; NID:6533205; PID:6533206
 #length 358 #molecular-weight 39760 #checksum 762
 SUMMARY
 Query Match 0.7% Score 7; DB 3; Length 358;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 245 PGNRRG 251
 QY 241 PGNRRG 247

RESULT 45
 #type complete
 #protein ubiquinol-cytochrome c reductase (EC 1.10.2.2) cytochrome b
 #formal_name (Saccharomyces cerevisiae) (common_name) man
 #date 14-Nov-1997 #sequence_revision 21-Jul-1995 #text_change
 #accessions 551087
 #references 551087
 #authors H. Kang, Y.W.; Jung, H.S.
 #journal Submitted to the EMBL Data Library, March 1995
 #title Sequence analysis of the cytochrome b gene of *Saccharomyces cerevisiae*
 #description papilionaceae mitochondria.
 #accession 551087
 #molecule_type 1-385 #label KIM
 #cross-reference EMBL:X8326; NID:6732979; PID:6732981
 #experimental_source strain CDC 75-737-D
 #genetics
 #gene cob
 #genome mitochondrion
 #genetic_code 111/3
 #status
 #classification superfamily cytochrome b; cytochrome b homology; cytochrome

CLASSIFICATION
 #superfamily galactokinase
 #EC 2.7.1.6
 #ATP_galactose_metalloprotein; phosphotransferase
 #length 385 #molecular-weight 40948 #checksum 5698
 SUMMARY
 Query Match 0.7% Score 7; DB 1; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 375 PGNRRG 381
 QY 260 PGNRRG 266

RESULT 47
 #type fragment
 #protein metacyclic form-specific variable surface glycoprotein B -
 #formal_name Trypanosoma congolense (strain IL1000) (fragment)
 #date 01-Jun-1993 #sequence_revision 09-Mar-1993 #text_change
 #accessions JCI142
 #references JCI141
 #authors Ehrlie, Y.; Urakawa, T.; Hirumi, H.; Flah, M.R.; Majima,
 Gene (1992) 113:139-146
 #journal Metacyclic form-specific variable surface
 #title glycoprotein encoding genes of *Trypanosoma* (Kinetoplastids)
 #cross-reference EMBL:9241662
 #accession JCI142
 #molecule_type mRNA
 #residues 1-425 #label EBR
 #length 425 #checksum 5443
 SUMMARY
 Query Match 0.7% Score 7; DB 2; Length 425;
 Best Local Similarity 100.0%; Pred. No. 1.04e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 270 AAFBALV 276
 QY 67 AAFBALV 72

RESULT 48
 #type complete
 #protein glucose-6-phosphate (G6P) dependent oxidoreductase STD - Streptomyces
 #formal_name Streptomyces glaucosens
 #date 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change
 #accession 555585
 #molecule_type DNA
 #genetics the source is designated as Streptomyces lividans
 #gene galk

ACCESSIONS 565585
 REFERENCE 565585
 AUTHOR Beyer, S.; Dieler, J.; Pieperberg, W.
 JOURNAL Mol. Gen. Genet. (1996) 256:75-78.
 TITLE The *hmd* gene of *Streptococcus glaucus* GLA.O
 (ETH 2734): new operon and evidence for pathway-specific
 regulation by STR.
 ACCESSION 565585 preliminary
 STATUS preliminary
 MOLECULE-type DNA
 STRESSES 1-428 #label BRY
 GENETICS across-reference EMBL: X89010; NID:980763; PID:980763
 SMODARY ATD
 length 428 molecular-weight 45633 checksum 6026
 QUERY MATCH 0.78; Score 7; DB 2; Length 428;
 Matched 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 259 DRYVGAH 275
 1111111
 659 DRYVGAH 655
 RESULT 49
 ENTRY 55335
 SMODARY 55335
 ORGANISM *Trichoderma harzianum*
 DATE 23-Aug-1995 sequence revision 19-Oct-1995 text-change
 19-Oct-1995
 ACCESSIONS 55335
 REFERENCE 55335
 AUTHOR Lore, J.M.; de la Cruz, J.; Lobell, A.; Benitez, T.;
 Pintor-Toro, J.A.
 JOURNAL Mol. Gen. Genet. (1995) 247:639-645
 TITLE Molecular cloning and characterization of an
 endo-beta-1,6-glucanase gene from the mycoparasitic fungus
Trichoderma harzianum.
 ACCESSION 55335
 STATUS preliminary
 MOLECULE-type 1-430 #label LOR
 SMODARY 55335
 length 430 molecular-weight 46165 checksum 6844
 QUERY MATCH 0.78; Score 7; DB 2; Length 430;
 Matched 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 160 ANYQDA 166

ACCESSIONS 569076
 REFERENCE 569076
 AUTHOR Dubois, J.; Alberge, T.; Bahlis, R.; Blakey, D.;
 Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.;
 Lum, W.; Pochter, B.; Qiu, D.; Spadator, R.; Vialle, R.;
 Wang, W.; Wierzbicki, J.; Gibson, R.; Jerni, N.; Caruso,
 McDonald, S.; Shiner, G.; Coyall, A.; Petrokovski, S.;
 Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Hoellings,
 J.; Reeve, J.N. (1997) 179:115-135
 JOURNAL Comp. Immun. Microbiol. (1997) 10:1-10
 TITLE Comparative genomics of *Methanobacterium*
thermautotrophicum Delta H: functional analysis and
 genome organization.
 ACCESSION 569076
 STATUS preliminary; nucleic acid sequence not shown;
 translation not shown
 MOLECULE-type DNA
 STRESSES 1-448 #label MTH
 GENETICS across-reference EMBL: X8907514
 SMODARY 569076
 length 448 molecular-weight 49096 checksum 4810
 QUERY MATCH 0.78; Score 7; DB 2; Length 448;
 Matched 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 56 88SLWNA 64
 817 88SLWNA 823
 Search completed: Fri Dec 18 18:50:22 1998
 Job time : 152 secs.

359	6	0.6	149	Y739_YEAST	HYPOPHOSPHATE 4.0 3 KD P	2.65
360	6	0.6	330	MC11_HUMAN	INDUCED MYELOID LEUKEM	2.65
361	6	0.6	350	ATP7B_HUMAN	HYDROPHASIS EXPRESSION	2.65
362	6	0.6	350	ATP7B_HUMAN	HYDROPHASIS EXPRESSION	2.65
363	6	0.6	351	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
364	6	0.6	351	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
365	6	0.6	352	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
366	6	0.6	352	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
367	6	0.6	353	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
368	6	0.6	353	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
369	6	0.6	353	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
370	6	0.6	353	HYFE_HUMAN	HYDROPHASIS EXPRESSION	2.65
371	6	0.6	357	SP11_YEAST	SP11_YEAST	2.65
372	6	0.6	357	SP11_YEAST	SP11_YEAST	2.65
373	6	0.6	358	ALP2_PLEA	ALP2_PLEA	2.65
374	6	0.6	358	ALP2_PLEA	ALP2_PLEA	2.65
375	6	0.6	359	ALP2_PLEA	ALP2_PLEA	2.65
376	6	0.6	359	ALP2_PLEA	ALP2_PLEA	2.65
377	6	0.6	359	ALP2_PLEA	ALP2_PLEA	2.65
378	6	0.6	360	DNH2_HUMAN	DNH2_HUMAN	2.65
379	6	0.6	360	DNH2_HUMAN	DNH2_HUMAN	2.65
380	6	0.6	360	DNH2_HUMAN	DNH2_HUMAN	2.65
381	6	0.6	360	DNH2_HUMAN	DNH2_HUMAN	2.65
382	6	0.6	361	F752_HUMAN	F752_HUMAN	2.65
383	6	0.6	361	F752_HUMAN	F752_HUMAN	2.65
384	6	0.6	362	ROG1_HUMAN	ROG1_HUMAN	2.65
385	6	0.6	362	ROG1_HUMAN	ROG1_HUMAN	2.65
386	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
387	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
388	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
389	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
390	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
391	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
392	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
393	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
394	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
395	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
396	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
397	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
398	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
399	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
400	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
401	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
402	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
403	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
404	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
405	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
406	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
407	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
408	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
409	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
410	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
411	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
412	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
413	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
414	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
415	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
416	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
417	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
418	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
419	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
420	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
421	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
422	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
423	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
424	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
425	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
426	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
427	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
428	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
429	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
430	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
431	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
432	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
433	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
434	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
435	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
436	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
437	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
438	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
439	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
440	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
441	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
442	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
443	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
444	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
445	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
446	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
447	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
448	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
449	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
450	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
451	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
452	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
453	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
454	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
455	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
456	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
457	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
458	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
459	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
460	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
461	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
462	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
463	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
464	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
465	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
466	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
467	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
468	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
469	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
470	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
471	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
472	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
473	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
474	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
475	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
476	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
477	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
478	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
479	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
480	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
481	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
482	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
483	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
484	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
485	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
486	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
487	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
488	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
489	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
490	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
491	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
492	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
493	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
494	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
495	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
496	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
497	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
498	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
499	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65
500	6	0.6	363	ADENYLATE CYCLAS 2 (E	ADENYLATE CYCLAS 2 (E	2.65

461	6	0.6	427	1.13	HUMAN	INTERLEUKIN-13 RECEPTOR
462	6	0.6	437	1.13	HUMAN	LEUCINE PROTEIN HOMO
463	6	0.6	439	1.13	YOUNG JAPANESE	HYDROLYTIC 46.9 KD P
464	6	0.6	439	1.13	YOUNG JAPANESE	COAT PROTEIN (PROTEIN
465	6	0.6	439	1.13	YOUNG JAPANESE	COAT PROTEIN (PROTEIN
466	6	0.6	430	1.13	MOUSE	PER B-CELL THERMOTA TR
467	6	0.6	430	1.13	MOUSE	OLIGOMERIN GLOSYLACT
468	6	0.6	430	1.13	MOUSE	NAODENDRYNIN TRANSPO
469	6	0.6	430	1.13	MOUSE	NAODENDRYNIN TRANSPO
470	6	0.6	430	1.13	MOUSE	NAODENDRYNIN TRANSPO
471	6	0.6	435	1.13	POBIA SPECIES	ADHEROSCAVING STRE
472	6	0.6	435	1.13	POBIA SPECIES	ADHEROSCAVING STRE
473	6	0.6	435	1.13	POBIA SPECIES	ADHEROSCAVING STRE
474	6	0.6	439	1.13	MOUSE	PROLACTIN RECEPTOR
475	6	0.6	439	1.13	MOUSE	PROLACTIN RECEPTOR
476	6	0.6	439	1.13	MOUSE	PROLACTIN RECEPTOR
477	6	0.6	439	1.13	MOUSE	PROLACTIN RECEPTOR
478	6	0.6	439	1.13	MOUSE	PROLACTIN RECEPTOR
479	6	0.6	440	1.13	MOUSE	PROLACTIN RECEPTOR
480	6	0.6	440	1.13	MOUSE	PROLACTIN RECEPTOR
481	6	0.6	441	1.13	MOUSE	PROLACTIN RECEPTOR
482	6	0.6	441	1.13	MOUSE	PROLACTIN RECEPTOR
483	6	0.6	441	1.13	MOUSE	PROLACTIN RECEPTOR
484	6	0.6	444	1.13	MOUSE	PROLACTIN RECEPTOR
485	6	0.6	444	1.13	MOUSE	PROLACTIN RECEPTOR
486	6	0.6	444	1.13	MOUSE	PROLACTIN RECEPTOR
487	6	0.6	444	1.13	MOUSE	PROLACTIN RECEPTOR
488	6	0.6	444	1.13	MOUSE	PROLACTIN RECEPTOR
489	6	0.6	446	1.13	MOUSE	PROLACTIN RECEPTOR
490	6	0.6	446	1.13	MOUSE	PROLACTIN RECEPTOR
491	6	0.6	446	1.13	MOUSE	PROLACTIN RECEPTOR
492	6	0.6	449	1.13	MOUSE	PROLACTIN RECEPTOR
493	6	0.6	451	1.13	MOUSE	PROLACTIN RECEPTOR
494	6	0.6	451	1.13	MOUSE	PROLACTIN RECEPTOR
495	6	0.6	452	1.13	MOUSE	PROLACTIN RECEPTOR
496	6	0.6	452	1.13	MOUSE	PROLACTIN RECEPTOR
497	6	0.6	453	1.13	MOUSE	PROLACTIN RECEPTOR
498	6	0.6	455	1.13	MOUSE	PROLACTIN RECEPTOR
499	6	0.6	455	1.13	MOUSE	PROLACTIN RECEPTOR
500	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
501	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
502	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
503	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
504	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
505	6	0.6	457	1.13	MOUSE	PROLACTIN RECEPTOR
506	6	0.6	459	1.13	MOUSE	PROLACTIN RECEPTOR
507	6	0.6	460	1.13	MOUSE	PROLACTIN RECEPTOR
508	6	0.6	460	1.13	MOUSE	PROLACTIN RECEPTOR
509	6	0.6	460	1.13	MOUSE	PROLACTIN RECEPTOR
510	6	0.6	461	1.13	MOUSE	PROLACTIN RECEPTOR
511	6	0.6	461	1.13	MOUSE	PROLACTIN RECEPTOR

410	6	0.6	381	1	C7B_MAOB	CYCLOCHROME B (EC 1.10.
411	6	0.6	381	1	C7B_MAOB	CYCLOCHROME B (EC 1.10.
412	6	0.6	381	1	C7B_MAOB	CYCLOCHROME B (EC 1.10.
413	6	0.6	381	1	C7B_MAOB	CYCLOCHROME B (EC 1.10.
414	6	0.6	382	1	N1P1_MOCB	HOMOCITRATE SYNTHASE (EC
415	6	0.6	382	1	S1U1_TGAB	PPE-MRNA SPLICING FACT
416	6	0.6	384	1	C7B_HSAF	CYCLOCHROME B (EC 1.10.
417	6	0.6	384	1	C7B_HSAF	CYCLOCHROME B (EC 1.10.
418	6	0.6	384	1	C7B_HSAF	CYCLOCHROME B (EC 1.10.
419	6	0.6	386	1	P3A1_HIDAN	PEPTIDIC ACID PHOSPHATASE
420	6	0.6	386	1	B1O1_AYTCT	ATLANTIC OXONADONATE 5
421	6	0.6	388	1	X1L1_STOUI	ATLANTIC ISOMERASE (EC 5
422	6	0.6	388	1	X1L1_STOUI	ATLANTIC ISOMERASE (EC 5
423	6	0.6	389	1	ATT1_SHIR	ATLANTIC AMINOXYTRANSFER
424	6	0.6	390	1	NMB1_HUDAN	NEUROKININ-B RECEPTOR
425	6	0.6	390	1	NMB1_HUDAN	NEUROKININ-B RECEPTOR
426	6	0.6	391	1	ENT1_SALAT	ETHANOLAMINE UTILITANT
427	6	0.6	391	1	ENT1_SALAT	ETHANOLAMINE UTILITANT
428	6	0.6	392	1	Y4P1_TGAB	HYDROLYTICAL 43.1.6 CD P
429	6	0.6	394	1	Y4L1_TGAB	RNA POLYMERASE PRINCIPLE
430	6	0.6	396	1	RND1_SATD	COLESTERE P430 HYDROGEN
431	6	0.6	398	1	P3H1_METVO	TRANSPOSSASE FOR INSERT
432	6	0.6	398	1	P3H1_METVO	TRANSPOSSASE FOR INSERT
433	6	0.6	399	1	T3A1_SHIR	PHOSPHOTRYPHOLITANT
434	6	0.6	399	1	P3R1_PBA	PHOSPHOTRYPHOLITANT
435	6	0.6	401	1	DT1_HAVLA	PHOSPHOTRYPHOLITANT
436	6	0.6	401	1	DT1_HAVLA	PHOSPHOTRYPHOLITANT
437	6	0.6	401	1	DT1_HAVLA	PHOSPHOTRYPHOLITANT
438	6	0.6	402	1	V4D1_PRAI	GLYCOPROTEIN GP50.
439	6	0.6	406	1	L4D1_MOOS	GLYCOPROTEIN ASSOCIATED
440	6	0.6	406	1	L4D1_MOOS	GLYCOPROTEIN ASSOCIATED
441	6	0.6	407	1	P3A1_P3PBU	3-CARBOXY-C18-C18-MO
442	6	0.6	407	1	L4P1_PAT	LEUCOSOME-ASSOCIATED ME
443	6	0.6	410	1	N1P1_MAOB	NEUTROPHIL KINASE (NKA)
444	6	0.6	410	1	N1P1_MAOB	NEUTROPHIL KINASE (NKA)
445	6	0.6	410	1	S1K1_AYTCT	SENSOR-LIKE HISTIDINE
446	6	0.6	410	1	OGB1_P3PBU	SENSOR-LIKE HISTIDINE
447	6	0.6	411	1	CG21_SCHPO	G2/ATOTIC SPECIFIC CT
448	6	0.6	411	1	CG21_SCHPO	G2/ATOTIC SPECIFIC CT
449	6	0.6	411	1	Y4S1_TGAB	HISTIDYL 47.8 ND P
450	6	0.6	411	1	Y4S1_TGAB	HISTIDYL 47.8 ND P
451	6	0.6	414	1	T3A1_STRO	MINI-CIRCLE HYDROLYTICAL
452	6	0.6	414	1	T3A1_STRO	MINI-CIRCLE HYDROLYTICAL
453	6	0.6	414	1	T3A1_STRO	MINI-CIRCLE HYDROLYTICAL
454	6	0.6	414	1	T3A1_STRO	MINI-CIRCLE HYDROLYTICAL
455	6	0.6	422	1	C7B1_AGAB	CYCLOCHROME P450-P4H1
456	6	0.6	424	1	C7B1_AGAB	CYCLOCHROME P450-P4H1
457	6	0.6	424	1	C7B1_AGAB	CYCLOCHROME P450-P4H1
458	6	0.6	424	1	C7B1_AGAB	CYCLOCHROME P450-P4H1
459	6	0.6	424	1	C7B1_AGAB	CYCLOCHROME P450-P4H1
460	6	0.6	426	1	C7B1_AGAB	CYCLOCHROME P450-P4H1

[illegible]

[illegible]

818	6	0.6	8311	1	CD44	HYDRA	CELL DIVISION CONTROL	2.6
819	6	0.6	835	1	CD44	YAK	CELL DIVISION CONTROL	2.6
820	6	0.6	835	1	CD44	YAK	CELL DIVISION CONTROL	2.6
821	6	0.6	848	1	CD44	YAK	CELL DIVISION CONTROL	2.6
822	6	0.6	848	1	CD44	YAK	CELL DIVISION CONTROL	2.6
823	6	0.6	849	1	CD44	YAK	CELL DIVISION CONTROL	2.6
824	6	0.6	852	1	CD44	YAK	CELL DIVISION CONTROL	2.6
825	6	0.6	852	1	CD44	YAK	CELL DIVISION CONTROL	2.6
826	6	0.6	862	1	CD44	YAK	CELL DIVISION CONTROL	2.6
827	6	0.6	863	1	CD44	YAK	CELL DIVISION CONTROL	2.6
828	6	0.6	870	1	CD44	YAK	CELL DIVISION CONTROL	2.6
829	6	0.6	870	1	CD44	YAK	CELL DIVISION CONTROL	2.6
830	6	0.6	882	1	CD44	YAK	CELL DIVISION CONTROL	2.6
831	6	0.6	882	1	CD44	YAK	CELL DIVISION CONTROL	2.6
832	6	0.6	882	1	CD44	YAK	CELL DIVISION CONTROL	2.6
833	6	0.6	882	1	CD44	YAK	CELL DIVISION CONTROL	2.6
834	6	0.6	882	1	CD44	YAK	CELL DIVISION CONTROL	2.6
835	6	0.6	883	1	CD44	YAK	CELL DIVISION CONTROL	2.6
836	6	0.6	883	1	CD44	YAK	CELL DIVISION CONTROL	2.6
837	6	0.6	883	1	CD44	YAK	CELL DIVISION CONTROL	2.6
838	6	0.6	883	1	CD44	YAK	CELL DIVISION CONTROL	2.6
839	6	0.6	883	1	CD44	YAK	CELL DIVISION CONTROL	2.6
840	6	0.6	892	1	CD44	YAK	CELL DIVISION CONTROL	2.6
841	6	0.6	915	1	CD44	YAK	CELL DIVISION CONTROL	2.6
842	6	0.6	915	1	CD44	YAK	CELL DIVISION CONTROL	2.6
843	6	0.6	915	1	CD44	YAK	CELL DIVISION CONTROL	2.6
844	6	0.6	911	1	CD44	YAK	CELL DIVISION CONTROL	2.6
845	6	0.6	918	1	CD44	YAK	CELL DIVISION CONTROL	2.6
846	6	0.6	918	1	CD44	YAK	CELL DIVISION CONTROL	2.6
847	6	0.6	918	1	CD44	YAK	CELL DIVISION CONTROL	2.6
848	6	0.6	920	1	CD44	YAK	CELL DIVISION CONTROL	2.6
849	6	0.6	923	1	CD44	YAK	CELL DIVISION CONTROL	2.6
850	6	0.6	924	1	CD44	YAK	CELL DIVISION CONTROL	2.6
851	6	0.6	924	1	CD44	YAK	CELL DIVISION CONTROL	2.6
852	6	0.6	924	1	CD44	YAK	CELL DIVISION CONTROL	2.6
853	6	0.6	930	1	CD44	YAK	CELL DIVISION CONTROL	2.6
854	6	0.6	932	1	CD44	YAK	CELL DIVISION CONTROL	2.6
855	6	0.6	950	1	CD44	YAK	CELL DIVISION CONTROL	2.6
856	6	0.6	950	1	CD44	YAK	CELL DIVISION CONTROL	2.6
857	6	0.6	952	1	CD44	YAK	CELL DIVISION CONTROL	2.6
858	6	0.6	952	1	CD44	YAK	CELL DIVISION CONTROL	2.6
859	6	0.6	953	1	CD44	YAK	CELL DIVISION CONTROL	2.6
860	6	0.6	961	1	CD44	YAK	CELL DIVISION CONTROL	2.6
861	6	0.6	963	1	CD44	YAK	CELL DIVISION CONTROL	2.6
862	6	0.6	977	1	CD44	YAK	CELL DIVISION CONTROL	2.6
863	6	0.6	979	1	CD44	YAK	CELL DIVISION CONTROL	2.6
864	6	0.6	979	1	CD44	YAK	CELL DIVISION CONTROL	2.6
865	6	0.6	982	1	CD44	YAK	CELL DIVISION CONTROL	2.6
866	6	0.6	982	1	CD44	YAK	CELL DIVISION CONTROL	2.6
867	6	0.6	993	1	CD44	YAK	CELL DIVISION CONTROL	2.6
868	6	0.6	993	1	CD44	YAK	CELL DIVISION CONTROL	2.6


```

CG      RIBH OR MYC2184 .34.
CC      HYPOCAERATION THYRUSTOSIS.
CC      HYPOCAERATION; FILICITES; ACTINOTCEGALUS; MICROBACTERIACEAE.
DN      [ ]
DN      SEQUENCE FROM N.A.
NA      KIRSH L., BARRETT D., BARRELL B.C., BALANDESKY M.A.;
NA      "HYPOCAERATION THYRUSTOSIS";
CC      -1- FUNCTION: RIBOFLAVIN SYNTHESIS IS A STRUCTURAL ENZYME COMPLEX
CC      CATALYZING THE FORMATION OF RIBOFLAVIN FROM 5-AMINO-6-(1'-D)-
CC      RIBITYL-AMINO-2-(4'II,3H) PYRIDIMEDIONE AND L-3,4-DIHYDROXY-2-
CC      BUTYL-CYCLOPENTANE-1,6-DIOL; DIHYDROXY-8-L-GLUTAMINE, THE BETA
CC      AMINO-2-(4'II,3H)-PYRIDIMEDIONE, AND L-3,4-DIHYDROXY-2-BUTANOXY-
CC      4-PHOSPHATE YIELDING 6,7-DIHYDRO-8-L-GLUTAMINE.
CC      -1- CATALYTIC ACTIVITY: 2,6,7-DIHYDRO-8-(1'-D-RIBITYL)GLUTAMINE +
CC      CYCLOPENTANE-1,6-DIOL = 5-AMINO-2,6-DIHYDROXYPYRIDINE.
CC      COFACTOR: FLAVOPROTEIN.
CC      -1- SUBUNIT: OLIGOMER THAT CONSIST OF 3 ALPHA SUBUNITS AND 60 BETA
CC      UNITS.
CC      ENCL: 280108; 2265577; -.
DR      RIBOFLAVIN BIOSYNTHESIS; TRANSFERASE; FLAVOPROTEIN.
SQ      SEQIDIVE 155 AA; 15904 NM; 9988086 CCM32.
Q? Query Match. 0.7% Score 7; DB 1; Length 154;
Best Local Similarity 100.0%; Pexd. No. 4,47e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0
DB      324 ALLOCAR 34
DY      324 ALLOCAR 130
|||||||
RESULT 16
AC      R47_S2HX          STANDARD:          PRT:    155 AA.
TC      R47_S2HX          33 CREATED)
DT      01-JUN-1995 (REV. 31, LAST ANNOTATION UPDATE)
DT      01-FEB-1996 (REV. 31, LAST ANNOTATION UPDATE)
DS      CHROMOPLAST 30S RIBOSOMAL PROTEIN S7.
DS      SPIROCYTIA MAXIMA.
OC      CHROMOPLAST.
OC      EUKARYOTA, PLANTA, PINCOPIPTA; CHLOROPIPTA (GREEN ALGA);
OC      CONJUGATOPHYTES; EUGLENIDALES; EUGENAZITHEALES; EUGENAZITACEAE.
RP      SEQUENCE FROM N.A.
RP      LBY K.A., BARRETT J.R.;
PJ      J. PROTOOL. 29:500-505(1993).

```

[illegible]

Mon Dec 21 09:28:24 1998

US-08-951-733-14.REF

Page 39

Mon Dec 21 09:28:24 1998

US-08-951-733-14.rtf

Page 40

[illegible]

```

FR 13 SEQUENCE FROM N.A. ARTICLE.
FR 14 MEDLINE: 89240011.
FR 15 KERNALIAN, J. & BODENMANN, R. 'TUMOR C.
FR 16 MCGILL ACTIN RES. 17:2751-2767(1989).
FR 17 SEQUENCE FROM X.A.
FR 18 STRAIVE-WISLA; TISSUE-HEART VENTRICLES;
FR 19 MEDLINE: 90016657.
FR 20 KERNALIAN, J. & BODENMANN, R. 'TUMOR C.
FR 21 MCGILL ACTIN RES. 17:2751-2767(1989).
FR 22 -1- SUBUNIT: MOGSI IS AN MEMBER OF 2 HEAVY CHAINS & 4 LIGHT CHAINS.
FR 23 -1- SIMILARITY TO OTHER EGF- AND CALCIUM BINDING PROTEINS, BUT THIS
FR 24 CC PROTEIN DOES NOT BIND CALCIUM.
FR 25 DB: X16325. 0763179..
FR 26 DB: X16326. 0763179.. JOINED.
FR 27 DB: X16327. 0763179.. JOINED.
FR 28 DB: X16328. 0763179.. JOINED.
FR 29 DB: X16329. 0763179.. JOINED.
FR 30 DB: PIR: 508573. 508573..
FR 31 DB: PIR: 506169. 506169..
FR 32 DB: PIR: 515759. 515759..
FR 33 DB: X16330. 0763179..
FR 34 DB: X16331. 0763179..
FR 35 DB: X16332. 0763179..
FR 36 DB: X16333. 0763179..
FR 37 DB: X16334. 0763179..
FR 38 DB: X16335. 0763179..
FR 39 DB: X16336. 0763179..
FR 40 DB: X16337. 0763179..
FR 41 DB: X16338. 0763179..
FR 42 DB: X16339. 0763179..
FR 43 DB: X16340. 0763179..
FR 44 DB: X16341. 0763179..
FR 45 DB: X16342. 0763179..
FR 46 DB: X16343. 0763179..
FR 47 DB: X16344. 0763179..
FR 48 DB: X16345. 0763179..
FR 49 DB: X16346. 0763179..
FR 50 DB: X16347. 0763179..
FR 51 DB: X16348. 0763179..
FR 52 DB: X16349. 0763179..
FR 53 DB: X16350. 0763179..
FR 54 DB: X16351. 0763179..
FR 55 DB: X16352. 0763179..
FR 56 DB: X16353. 0763179..
FR 57 DB: X16354. 0763179..
FR 58 DB: X16355. 0763179..
FR 59 DB: X16356. 0763179..
FR 60 DB: X16357. 0763179..
FR 61 DB: X16358. 0763179..
FR 62 DB: X16359. 0763179..
FR 63 DB: X16360. 0763179..
FR 64 DB: X16361. 0763179..
FR 65 DB: X16362. 0763179..
FR 66 DB: X16363. 0763179..
FR 67 DB: X16364. 0763179..
FR 68 DB: X16365. 0763179..
FR 69 DB: X16366. 0763179..
FR 70 DB: X16367. 0763179..
FR 71 DB: X16368. 0763179..
FR 72 DB: X16369. 0763179..
FR 73 DB: X16370. 0763179..
FR 74 DB: X16371. 0763179..
FR 75 DB: X16372. 0763179..
FR 76 DB: X16373. 0763179..
FR 77 DB: X16374. 0763179..
FR 78 DB: X16375. 0763179..
FR 79 DB: X16376. 0763179..
FR 80 DB: X16377. 0763179..
FR 81 DB: X16378. 0763179..
FR 82 DB: X16379. 0763179..
FR 83 DB: X16380. 0763179..
FR 84 DB: X16381. 0763179..
FR 85 DB: X16382. 0763179..
FR 86 DB: X16383. 0763179..
FR 87 DB: X16384. 0763179..
FR 88 DB: X16385. 0763179..
FR 89 DB: X16386. 0763179..
FR 90 DB: X16387. 0763179..
FR 91 DB: X16388. 0763179..
FR 92 DB: X16389. 0763179..
FR 93 DB: X16390. 0763179..
FR 94 DB: X16391. 0763179..
FR 95 DB: X16392. 0763179..
FR 96 DB: X16393. 0763179..
FR 97 DB: X16394. 0763179..
FR 98 DB: X16395. 0763179..
FR 99 DB: X16396. 0763179..
FR 100 DB: X16397. 0763179..
FR 101 DB: X16398. 0763179..
FR 102 DB: X16399. 0763179..
FR 103 DB: X16400. 0763179..
FR 104 DB: X16401. 0763179..
FR 105 DB: X16402. 0763179..
FR 106 DB: X16403. 0763179..
FR 107 DB: X16404. 0763179..
FR 108 DB: X16405. 0763179..
FR 109 DB: X16406. 0763179..
FR 110 DB: X16407. 0763179..
FR 111 DB: X16408. 0763179..
FR 112 DB: X16409. 0763179..
FR 113 DB: X16410. 0763179..
FR 114 DB: X16411. 0763179..
FR 115 DB: X16412. 0763179..
FR 116 DB: X16413. 0763179..
FR 117 DB: X16414. 0763179..
FR 118 DB: X16415. 0763179..
FR 119 DB: X16416. 0763179..
FR 120 DB: X16417. 0763179..
FR 121 DB: X16418. 0763179..
FR 122 DB: X16419. 0763179..
FR 123 DB: X16420. 0763179..
FR 124 DB: X16421. 0763179..
FR 125 DB: X16422. 0763179..
FR 126 DB: X16423. 0763179..
FR 127 DB: X16424. 0763179..
FR 128 DB: X16425. 0763179..
FR 129 DB: X16426. 0763179..
FR 130 DB: X16427. 0763179..
FR 131 DB: X16428. 0763179..
FR 132 DB: X16429. 0763179..
FR 133 DB: X16430. 0763179..
FR 134 DB: X16431. 0763179..
FR 135 DB: X16432. 0763179..
FR 136 DB: X16433. 0763179..
FR 137 DB: X16434. 0763179..
FR 138 DB: X16435. 0763179..
FR 139 DB: X16436. 0763179..
FR 140 DB: X16437. 0763179..
FR 141 DB: X16438. 0763179..
FR 142 DB: X16439. 0763179..
FR 143 DB: X16440. 0763179..
FR 144 DB: X16441. 0763179..
FR 145 DB: X16442. 0763179..
FR 146 DB: X16443. 0763179..
FR 147 DB: X16444. 0763179..
FR 148 DB: X16445. 0763179..
FR 149 DB: X16446. 0763179..
FR 150 DB: X16447. 0763179..
FR 151 DB: X16448. 0763179..
FR 152 DB: X16449. 0763179..
FR 153 DB: X16450. 0763179..
FR 154 DB: X16451. 0763179..
FR 155 DB: X16452. 0763179..
FR 156 DB: X16453. 0763179..
FR 157 DB: X16454. 0763179..
FR 158 DB: X16455. 0763179..
FR 159 DB: X16456. 0763179..
FR 160 DB: X16457. 0763179..
FR 161 DB: X16458. 0763179..
FR 162 DB: X16459. 0763179..
FR 163 DB: X16460. 0763179..
FR 164 DB: X16461. 0763179..
FR 165 DB: X16462. 0763179..
FR 166 DB: X16463. 0763179..
FR 167 DB: X16464. 0763179..
FR 168 DB: X16465. 0763179..
FR 169 DB: X16466. 0763179..
FR 170 DB: X16467. 0763179..
FR 171 DB: X16468. 0763179..
FR 172 DB: X16469. 0763179..
FR 173 DB: X16470. 0763179..
FR 174 DB: X16471. 0763179..
FR 175 DB: X16472. 0763179..
FR 176 DB: X16473. 0763179..
FR 177 DB: X16474. 0763179..
FR 178 DB: X16475. 0763179..
FR 179 DB: X16476. 0763179..
FR 180 DB: X16477. 0763179..
FR 181 DB: X16478. 0763179..
FR 182 DB: X16479. 0763179..
FR 183 DB: X16480. 0763179..
FR 184 DB: X16481. 0763179..
FR 185 DB: X16482. 0763179..
FR
```

```

RA MARKER:SON S.;
RL SOLGATED (INV-1985) TO EMBL/GENBANK/DDBB DATA BANKS.
CC CLEAVAGE OF THE PROTEIN AT THE SITE OF THE NUCLEOTIDE SEQUENCE.
DR EMBL, EMBL055, G1073446; - PROTEIN 592;.
DR NONREPEATED: P1366.1; CRO1398;.
RV HYPOPHYSICAL PROTEIN.
SV SEQUENCE 257 AA: 26656 MW; 66FD9900 CRC32;.
Q Query Match 0.7%; Score 7; DB 3; Length 257;
Best Local Similarity: 100.0%; Pos: No. 4,47e+00;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;.
Db 30 DAWVIDO 56
Qy 810 DAWVIDO 816

RESULT 21
ID BIA:AD07 STANDARD: PRT: 261 AA.
DC C POLYMERASE-1866 (REL. 01. CHARGE);
D7 21-NOV-1986 (REL. 01. LAST SEQUENCE UPDATE);
D7 01-NOV-1997 (REL. 35; LAST AMENDMENT UPDATE);
D7 BAUKA VIA 28 DN NONDEVELOPED VIRUSES; EARLY ELA 24 CD PROTEIN);
OC VIRALIA; DS-DNA NONDEVELOPED VIRUSES; ADENOVIRALINE; NASUTUDONVIRUSIS.
RM [1];
RZ SEQUENCE FROM R.A.A.
RC STRAIN:COHEN;
RA DICTIONARY: 71792
RA BOVEN H.W.; DEKKE B.M.H.; VAN OORDNET H.; DE WAND A.; BLAT J.;
RL Gene 12:287-299(1980).
RZ SEQUENCE FROM R.A.A.
RC STRAIN:GIDER;
RA TOSHIDA K.; FUJIMURA K.;
RL YOSHIKAWA K.; 35-37(1984);
CC -1- PROMOTING FACTORS ACTIVATES EARLY VIRAL PROMOTERS AND SOME CELLULAR
CC NO PROTEIN ARE PRODUCTS OF THESE DIFFERENT REGIONS SPICED FROM THE
CC PROMOTING FACTORS ACTIVATES EARLY VIRAL PROMOTERS AND SOME CELLULAR
CC NO PROTEIN ARE PRODUCTS OF THESE DIFFERENT REGIONS SPICED FROM THE
DR EMBL, G104000; G58515;.
DR EMBL, X03000; G58517;.
DR EMBL, M18648; G209784;.
DR EMBL, M18648; G209784;.
DR EMBL, M18648; G209784;.
RV TRANSCRIPTION REGULATOR: ACTIVATOR; EARLY PROTEIN;
RV TRANSCRIPTION REGULATOR: ZINC-FINGER; DNA-BINDING.
RV ALTERNATIVE SPLICING: ZINC-FINGER; DNA-BINDING.
CL TYPE.
RL 169 183
T

```

```

FT VASISALIC 164 194 MISSING (IN ENAL ELA 24 KD PROTEIN).
FT VARIANT 164 262 C->E (IN STRAIN GRIDER).
FT VARIANT 202 262 P (IN STRAIN GRIDER).
FT SEQUENCE 261 AA, 28355 MW, 724362Da (CH27).
Query Match 0.7% Score 7: Db 1: Length 261.
Best Local Similarity 100.0% Prd. No. 4,47e-004
Matches 7: Conservative 0: Miscellaneous 0: Indels 0: Gaps 0
Db 178 LIGSCLCT 184
O7 862 LIGSCLCT 868
|||||||
RESULT 32
1D 1D TRP, IRMAN STANDARD: PRT: 378 AA.
1E 1E
1F 1F
1G 1G
1H 1H
1I 1I
1J 1J
1K 1K
1L 1L
1M 1M
1N 1N
1O 1O
1P 1P
1Q 1Q
1R 1R
1S 1S
1T 1T
1U 1U
1V 1V
1W 1W
1X 1X
1Y 1Y
1Z 1Z
2D 2D
2E 2E
2F 2F
2G 2G
2H 2H
2I 2I
2J 2J
2K 2K
2L 2L
2M 2M
2N 2N
2O 2O
2P 2P
2Q 2Q
2R 2R
2S 2S
2T 2T
2U 2U
2V 2V
2W 2W
2X 2X
2Y 2Y
2Z 2Z
3D 3D
3E 3E
3F 3F
3G 3G
3H 3H
3I 3I
3J 3J
3K 3K
3L 3L
3M 3M
3N 3N
3O 3O
3P 3P
3Q 3Q
3R 3R
3S 3S
3T 3T
3U 3U
3V 3V
3W 3W
3X 3X
3Y 3Y
3Z 3Z
4D 4D
4E 4E
4F 4F
4G 4G
4H 4H
4I 4I
4J 4J
4K 4K
4L 4L
4M 4M
4N 4N
4O 4O
4P 4P
4Q 4Q
4R 4R
4S 4S
4T 4T
4U 4U
4V 4V
4W 4W
4X 4X
4Y 4Y
4Z 4Z
5D 5D
5E 5E
5F 5F
5G 5G
5H 5H
5I 5I
5J 5J
5K 5K
5L 5L
5M 5M
5N 5N
5O 5O
5P 5P
5Q 5Q
5R 5R
5S 5S
5T 5T
5U 5U
5V 5V
5W 5W
5X 5X
5Y 5Y
5Z 5Z
6D 6D
6E 6E
6F 6F
6G 6G
6H 6H
6I 6I
6J 6J
6K 6K
6L 6L
6M 6M
6N 6N
6O 6O
6P 6P
6Q 6Q
6R 6R
6S 6S
6T 6T
6U 6U
6V 6V
6W 6W
6X 6X
6Y 6Y
6Z 6Z
7D 7D
7E 7E
7F 7F
7G 7G
7H 7H
7I 7I
7J 7J
7K 7K
7L 7L
7M 7M
7N 7N
7O 7O
7P 7P
7Q 7Q
7R 7R
7S 7S
7T 7T
7U 7U
7V 7V
7W 7W
7X 7X
7Y 7Y
7Z 7Z
8D 8D
8E 8E
8F 8F
8G 8G
8H 8H
8I 8I
8J 8J
8K 8K
8L 8L
8M 8M
8N 8N
8O 8O
8P 8P
8Q 8Q
8R 8R
8S 8S
8T 8T
8U 8U
8V 8V
8W 8W
8X 8X
8Y 8Y
8Z 8Z
9D 9D
9E 9E
9F 9F
9G 9G
9H 9H
9I 9I
9J 9J
9K 9K
9L 9L
9M 9M
9N 9N
9O 9O
9P 9P
9Q 9Q
9R 9R
9S 9S
9T 9T
9U 9U
9V 9V
9W 9W
9X 9X
9Y 9Y
9Z 9Z
10D 10D
10E 10E
10F 10F
10G 10G
10H 10H
10I 10I
10J 10J
10K 10K
10L 10L
10M 10M
10N 10N
10O 10O
10P 10P
10Q 10Q
10R 10R
10S 10S
10T 10T
10U 10U
10V 10V
10W 10W
10X 10X
10Y 10Y
10Z 10Z
11D 11D
11E 11E
11F 11F
11G 11G
11H 11H
11I 11I
11J 11J
11K 11K
11L 11L
11M 11M
11N 11N
11O 11O
11P 11P
11Q 11Q
11R 11R
11S 11S
11T 11T
11U 11U
11V 11V
11W 11W
11X 11X
11Y 11Y
11Z 11Z
12D 12D
12E 12E
12F 12F
12G 12G
12H 12H
12I 12I
12J 12J
12K 12K
12L 12L
12M 12M
12N 12N
12O 12O
12P 12P
12Q 12Q
12R 12R
12S 12S
12T 12T
12U 12U
12V 12V
12W 12W
12X 12X
12Y 12Y
12Z 12Z
13D 13D
13E 13E
13F 13F
13G 13G
13H 13H
13I 13I
13J 13J
13K 13K
13L 13L
13M 13M
13N 13N
13O 13O
13P 13P
13Q 13Q
13R 13R
13S 13S
13T 13T
13U 13U
13V 13V
13W 13W
13X 13X
13Y 13Y
13Z 13Z
14D 14D
14E 14E
14F 14F
14G 14G
14H 14H
14I 14I
14J 14J
14K 14K
14L 14L
14M 14M
14N 14N
14O 14O
14P 14P
14Q 14Q
14R 14R
14S 14S
14T 14T
14U 14U
14V 14V
14W 14W
14X 14X
14Y 14Y
14Z 14Z
15D 15D
15E 15E
15F 15F
15G 15G
15H 15H
15I 15I
15J 15J
15K 15K
15L 15L
15M 15M
15N 15N
15O 15O
15P 15P
15Q 15Q
15R 15R
15S 15S
15T 15T
15U 15U
15V 15V
15W 15W
15X 15X
15Y 15Y
15Z 15Z
16D 16D
16E 16E
16F 16F
16G 16G
16H 16H
16I 16I
16J 16J
16K 16K
16L 16L
16M 16M
16N 16N
16O 16O
16P 16P
16Q 16Q
16R 16R
16S 16S
16T 16T
16U 16U
16V 16V
16W 16W
16X 16X
16Y 16Y
16Z 16Z
17D 17D
17E 17E
17F 17F
17G 17G
17H 17H
17I 17I
17J 17J
17K 17K
17L 17L
17M 17M
17N 17N
17O 17O
17P 17P
17Q 17Q
17R 17R
17S 17S
17T 17T
17U 17U
17V 17V
17W 17W
17X 17X
17Y 17Y
17Z 17Z
18D 18D
18E 18E
18F 18F
18G 18G
18H 18H
18I 18I
18J 18J
18K 18K
18L 18L
18M 18M
18N 18N
18O 18O
18P 18P
18Q 18Q
18R 18R
18S 18S
18T 18T
18U 18U
18V 18V
18W 18W
18X 18X
18Y 18Y
18Z 18Z
19D 19D
19E 19E
19F 19F
19G 19G
19H 19H
19I 19I
19J 19J
19K 19K
19L 19L
19M 19M
19N 19N
19O 19O
19P 19P
19Q 19Q
19R 19R
19S 19S
19T 19T
19U 19U
19V 19V
19W 19W
19X 19X
19Y 19Y
19Z 19Z
20D 20D
20E 20E
20F 20F
20G 20G
20H 20H
20I 20I
20J 20J
20K 20K
20L 20L
20M 20M
20N 20N
20O 20O
20P 20P
20Q 20Q
20R 20R
20S 20S
20T 20T
20U 20U
20V 20V
20W 20W
20X 20X
20Y 20Y
20Z 20Z
21D 21D
21E 21E
21F 21F
21G 21G
21H 21H
21I 21I
21J 21J
21K 21K
21L 21L
21M 21M
21N 21N
21O 21O
21P 21P
21Q 21Q
21R 21R
21S 21S
21T 21T
21U 21U
21V 21V
21W 21W
21X 21X
21Y 21Y
21Z 21Z
22D 22D
22E 22E
22F 22F
22G 22G
22H 22H
22I 22I
22J 22J
22K 22K
22L 22L
22M 22M
22N 22N
22O 22O
22P 22P
22Q 22Q
22R 22R
22S 22S
22T 22T
22U 22U
22V 22V
22W 22W
22X 22X
22Y 22Y
22Z 22Z
23D 23D
23E 23E
23F 23F
23G 23G
23H 23H
23I 23I
23J 23J
23K 23K
23L 23L
23M 23M
23N 23N
23O 23O
23P 23P
23Q 23Q
23R 23R
23S 23S
23T 23T
23U 23U
23V 23V
23W 23W
23X 23X
23Y 23Y
23Z 23Z
24D 24D
24E 24E
24F 24F
24G 24G
24H 24H
24I 2
```

[illegible][illegible]

01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE PEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN SAPB.
 DE SAMPONELLA TYPHIMURIM.
 OC SAMPONELLA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;
 CC ESTEROBACTERIACEAE.
 RC STRAIN-140285.
 RX MEDLINE: 94038887.
 BA PARA-LOPEZ C.; BAKER M.T.; GROISMAN E.A.;
 RL EMBL: X74217; G414209; -.
 CC -1- FUNCTION: TRANSPORT. PEPTIDE INTRATE TRANSPORT SYSTEM THAT
 CC -1- PLAYS A ROLE IN THE RESISTANCE TO ANTIBIOTIC PEPTIDES
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -1- (POTENTIAL). WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC -1- PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPRC
 CC SUBFAMILY.
 DR EMBL: X74217; G414209; -.
 DR PIR: S35566; S35568.
 DR PROSITE: PS00402; BPD_TRANSP_INN_MEMB; PALSE_NRG.
 DE TRANSPORT. PEPTIDE TRANSPORT. TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 321 AA; 36093 MW; 6531780 CIRC32;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 4,47e+00; Length 321;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 83 LARGAL 89
 119 LARGAL 125
 119 LARGAL 125

01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE PEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN SAPB.
 DE SAMPONELLA TYPHIMURIM.
 OC SAMPONELLA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;
 CC ESTEROBACTERIACEAE.
 RC STRAIN-140285.
 RX MEDLINE: 94038887.
 BA PARA-LOPEZ C.; BAKER M.T.; GROISMAN E.A.;
 RL EMBL: X74217; G414209; -.
 CC -1- FUNCTION: TRANSPORT. PEPTIDE INTRATE TRANSPORT SYSTEM THAT
 CC -1- PLAYS A ROLE IN THE RESISTANCE TO ANTIBIOTIC PEPTIDES
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -1- (POTENTIAL). WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC -1- PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPRC
 CC SUBFAMILY.
 DR EMBL: X74217; G414209; -.
 DR PIR: S35566; S35568.
 DR PROSITE: PS00402; BPD_TRANSP_INN_MEMB; PALSE_NRG.
 DE TRANSPORT. PEPTIDE TRANSPORT. TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 321 AA; 36093 MW; 6531780 CIRC32;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 4,47e+00; Length 321;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 83 LARGAL 89
 119 LARGAL 125
 119 LARGAL 125

01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE PEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN SAPB.
 DE SAMPONELLA TYPHIMURIM.
 OC SAMPONELLA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;
 CC ESTEROBACTERIACEAE.
 RC STRAIN-140285.
 RX MEDLINE: 94038887.
 BA PARA-LOPEZ C.; BAKER M.T.; GROISMAN E.A.;
 RL EMBL: X74217; G414209; -.
 CC -1- FUNCTION: TRANSPORT. PEPTIDE INTRATE TRANSPORT SYSTEM THAT
 CC -1- PLAYS A ROLE IN THE RESISTANCE TO ANTIBIOTIC PEPTIDES
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -1- (POTENTIAL). WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC -1- PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPRC
 CC SUBFAMILY.
 DR EMBL: X74217; G414209; -.
 DR PIR: S35566; S35568.
 DR PROSITE: PS00402; BPD_TRANSP_INN_MEMB; PALSE_NRG.
 DE TRANSPORT. PEPTIDE TRANSPORT. TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 321 AA; 36093 MW; 6531780 CIRC32;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 4,47e+00; Length 321;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 83 LARGAL 89
 119 LARGAL 125
 119 LARGAL 125

01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE PEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN SAPB.
 DE SAMPONELLA TYPHIMURIM.
 OC SAMPONELLA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;
 CC ESTEROBACTERIACEAE.
 RC STRAIN-140285.
 RX MEDLINE: 94038887.
 BA PARA-LOPEZ C.; BAKER M.T.; GROISMAN E.A.;
 RL EMBL: X74217; G414209; -.
 CC -1- FUNCTION: TRANSPORT. PEPTIDE INTRATE TRANSPORT SYSTEM THAT
 CC -1- PLAYS A ROLE IN THE RESISTANCE TO ANTIBIOTIC PEPTIDES
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -1- (POTENTIAL). WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC -1- PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPRC
 CC SUBFAMILY.
 DR EMBL: X74217; G414209; -.
 DR PIR: S35566; S35568.
 DR PROSITE: PS00402; BPD_TRANSP_INN_MEMB; PALSE_NRG.
 DE TRANSPORT. PEPTIDE TRANSPORT. TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 321 AA; 36093 MW; 6531780 CIRC32;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 4,47e+00; Length 321;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 287 ASRSUPL 293
 250 ASRSUPL 256
 250 ASRSUPL 256

01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE PEPTIDE TRANSPORT SYSTEM PERMEASE PROTEIN SAPB.
 DE SAMPONELLA TYPHIMURIM.
 OC SAMPONELLA: GRACILICUTES; SCOTOBACTERIA: FACULTATIVELY ANAEROBIC RODS;
 CC ESTEROBACTERIACEAE.
 RC STRAIN-140285.
 RX MEDLINE: 94038887.
 BA PARA-LOPEZ C.; BAKER M.T.; GROISMAN E.A.;
 RL EMBL: X74217; G414209; -.
 CC -1- FUNCTION: TRANSPORT. PEPTIDE INTRATE TRANSPORT SYSTEM THAT
 CC -1- PLAYS A ROLE IN THE RESISTANCE TO ANTIBIOTIC PEPTIDES
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
 CC -1- (POTENTIAL). WITH INTEGRAL MEMBRANE COMPONENTS OF OTHER BINDING-
 CC -1- PROTEIN-DEPENDENT TRANSPORT SYSTEMS. BELONGS TO THE OPRC
 CC SUBFAMILY.
 DR EMBL: X74217; G414209; -.
 DR PIR: S35566; S35568.
 DR PROSITE: PS00402; BPD_TRANSP_INN_MEMB; PALSE_NRG.
 DE TRANSPORT. PEPTIDE TRANSPORT. TRANSMEMBRANE; INNER MEMBRANE.
 FT TRANSMEM 81 101 POTENTIAL.
 FT TRANSMEM 114 134 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 249 269 POTENTIAL.
 FT TRANSMEM 281 301 POTENTIAL.
 SQ SEQUENCE 321 AA; 36093 MW; 6531780 CIRC32;
 Query Match
 Best Local Similarity 100.0%; Pred. No. 4,47e+00; Length 321;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 287 ASRSUPL 293
 250 ASRSUPL 256
 250 ASRSUPL 256

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

164	6	0	6	723.13	0.93917	ACTIN FILAMENT PROTEIN
165	6	0	6	727.14	0.61306	ISOLATE CPV-126 CONT P
166	6	0	6	727.14	0.61306	ISOLATE CPV-126 CONT P
167	6	0	6	727.14	0.61306	ISOLATE CPV-126 CONT P
168	6	0	6	729.2	0.67823	HYDROLYTIC 74.5 KD P
169	6	0	6	729.21	0.61827	PHOSPHATASE GROWTH PACT
170	6	0	6	729.21	0.61827	MAJOR CONT PROTEIN VAI
171	6	0	6	735.14	0.56524	CONT PROTEIN VAI
172	6	0	6	737.14	0.90486	CASID PROTEIN VAI
173	6	0	6	737.14	0.90486	CASID PROTEIN VAI
174	6	0	6	737.14	0.90486	CASID PROTEIN VAI
175	6	0	6	737.14	0.90486	CASID PROTEIN VAI
176	6	0	6	739.4	0.41554	ATP-BINDING
177	6	0	6	745.4	0.15151	PHOSPHOLIPASE C (EC 3.
178	6	0	6	745.6	0.329517	PHOSPHOLIPASE C (EC 3.
179	6	0	6	746.4	0.12586	BRUSH BORDER TRANSPORT
180	6	0	6	746.5	0.27556	BETA-CALCOTRIOLINASE (EC
181	6	0	6	746.5	0.27556	HYDROLYTIC 74.5 KD P
182	6	0	6	747.5	0.27476	HYDROLYTIC 74.5 KD P
183	6	0	6	749.4	0.13816	PROTEIN
184	6	0	6	749.4	0.13816	PROTEIN
185	6	0	6	749.4	0.13816	PROTEIN
186	6	0	6	749.4	0.13816	PROTEIN
187	6	0	6	749.4	0.13816	PROTEIN
188	6	0	6	749.4	0.13816	PROTEIN
189	6	0	6	749.4	0.13816	PROTEIN
190	6	0	6	751.5	0.01457	SHIMILAR TO THE THROMBIN
191	6	0	6	752.10	0.04319	THROMBIN
192	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
193	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
194	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
195	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
196	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
197	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
198	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
199	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
200	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
201	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
202	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
203	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
204	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
205	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
206	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
207	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
208	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
209	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
210	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
211	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
212	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
213	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT
214	6	0	6	752.2	0.00592	D-(-)-3-HYDROXYBUTYRAT

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

866	6	0	6	974.5	892165	2K710_ZC (FRANCERY)
867	6	0	6	975.1	910728	SENTRIN/UBIQUITIN-PROTE
868	6	0	6	975.1	910728	SENTRIN/UBIQUITIN-PROTE
869	6	0	6	995.1	809658	TAIL-SPECIFIC THROMBIN
870	6	0	6	999.11	064518	SACCHAROPLASMIC RETIC
871	6	0	6	1002.11	0153596	TRANSCRIPTION FACTOR
872	6	0	6	1002.11	0153596	TRANSCRIPTION FACTOR
873	6	0	6	1005.4	024154	HELIX-LOOP-HELIX TRANS
874	6	0	6	1037.5	024155	TRANSCRIPTION FACTOR
875	6	0	6	1038.11	064511	SACCHAROPLASMIC RETIC
876	6	0	6	1040.11	090840	A VARIANT OF TSC-22
877	6	0	6	1052.5	026518	KITOPLAST-ASSOCIATED
878	6	0	6	1053.1	065393	POLY(POLYMERIN)
879	6	0	6	1055.2	0153148	HYDROLYTICAL, 110.7 KD
880	6	0	6	1058.5	029505	BIOTINYL-CATALANIN
881	6	0	6	1058.5	029505	BIOTINYL-CATALANIN
882	6	0	6	1066.1	029505	BIOTINYL-CATALANIN
883	6	0	6	1066.1	029505	BIOTINYL-CATALANIN
884	6	0	6	1068.1	0242391	HEPES SIMPLEX VARIANTS
885	6	0	6	1080.2	048379	PHOSPHOINOSITIDE 3-KIN
886	6	0	6	1080.2	048379	PHOSPHOINOSITIDE 3-KIN
887	6	0	6	1121.1	0109862	ADENOSINE 120.6 KD
888	6	0	6	1126.5	013971	PLATELET
889	6	0	6	1130.1	088282	ENZYMOLASE PROTEIN
890	6	0	6	1133.4	016514	ZINC METALLOPROTEIN
891	6	0	6	1133.4	016514	ZINC METALLOPROTEIN
892	6	0	6	1182.5	024818	HYDROLYTICAL, 177.3 KD
893	6	0	6	1185.2	0295031	HYDROLYTICAL, 177.3 KD
894	6	0	6	1175.11	015126	DETA
895	6	0	6	1189.5	017673	COSEY 5
896	6	0	6	1189.5	017673	COSEY 5
897	6	0	6	1181.2	052484	WEIR
898	6	0	6	1194.2	053445	PURINIC ABC TRANSPORT
899	6	0	6	1205.1	010846	NITRIC OXIDE REDUCTASE
900	6	0	6	1218.5	023126	SIMILAR TO MAP KINASE
901	6	0	6	1219.2	023185	TRANSMEMBRANE PROTEIN
902	6	0	6	1222.5	029692	RELATIVES TO PH-A17124
903	6	0	6	1262.5	046639	EC 6B5.2 PROTEIN
904	6	0	6	1265.5	024850	P-GLYCOPROTEIN 5
905	6	0	6	1301.5	024850	P-GLYCOPROTEIN 5
906	6	0	6	1304.5	015765	CYTOSOLIC DYKIN H2A
907	6	0	6	1304.5	015765	CYTOSOLIC DYKIN H2A
908	6	0	6	1344.1	008851	POLYMERIN (PARENTS)
909	6	0	6	1344.1	008851	POLYMERIN (PARENTS)
910	6	0	6	1359.10	023037	ALPHA 2 TYPE I COLLAGEN
911	6	0	6	1359.10	023037	ALPHA 2 TYPE I COLLAGEN
912	6	0	6	1375.1	029381	COMPLEMENT OF BRYN 1
913	6	0	6	1375.1	029381	COMPLEMENT OF BRYN 1
914	6	0	6	1375.1	029381	COMPLEMENT OF BRYN 1
915	6	0	6	1375.1	029381	COMPLEMENT OF BRYN 1
916	6	0	6	1375.11	035059	ADPOPHOS. SIGNAL-RECU

917	6	0.6	1408.3	Q43140	WILDTYPE 5'-UTRANSLAC	5
918	6	0.6	1408.3	Q43140	WILDTYPE 5'-UTRANSLAC	5
919	6	0.6	1428.4	Q50574	KIA0366 PROTEIN P13C	5
920	6	0.6	1441.5	Q26453	HIS-HOMOLESS (FRAGMENT	5
921	6	0.6	1442.1	Q38358	RNA POLYMERASE	5
922	6	0.6	1463.1	Q54446	COPPER TRANSPORTING P-	5
923	6	0.6	1463.1	Q54446	COPPER TRANSPORTING P-	5
924	6	0.6	1525.5	Q34328	PALELLED	5
925	6	0.6	1544.1	Q53489	MITOCHONDRIAL KIA0359 (P	5
926	6	0.6	1550.4	Q52487	SULFONOLACTAM RECEPTOR	5
927	6	0.6	1558.4	Q43434	RNA 1.1.1.1.1.1.1.1.1.1.1	5
928	6	0.6	1625.1	Q55597	RNA 1.1.1.1.1.1.1.1.1.1.1	5
929	6	0.6	1625.1	Q55597	RNA 1.1.1.1.1.1.1.1.1.1.1	5
930	6	0.6	1646.14	Q86820	POLYMERASE	5
931	6	0.6	1646.14	Q86820	POLYMERASE	5
932	6	0.6	1646.14	Q86820	POLYMERASE	5
933	6	0.6	1646.14	Q86820	POLYMERASE	5
934	6	0.6	1667.11	Q52324	SHEDDING	5
935	6	0.6	1677.5	Q00805	SHEDDING	5
936	6	0.6	1724.5	Q53425	P-GLYCOPROTEIN E.	5
937	6	0.6	1724.5	Q53425	P-GLYCOPROTEIN E.	5
938	6	0.6	1724.5	Q53425	P-GLYCOPROTEIN E.	5
939	6	0.6	1766.11	Q53546	POLYMERASE	5
940	6	0.6	1765.5	Q54385	LATELY TOP-BETA BINDING	5
941	6	0.6	1803.3	Q00090	MAJOR MEMBRANE SURFACE	5
942	6	0.6	1803.3	Q00090	MAJOR MEMBRANE SURFACE	5
943	6	0.6	1811.14	Q51684	POLYMERASE	5
944	6	0.6	1841.4	Q15031	NONSTRUCTURAL POLYPEPT	5
945	6	0.6	1841.4	Q15031	NONSTRUCTURAL POLYPEPT	5
946	6	0.6	1846.5	Q58121	KIA0015 (FRAGMENT)	5
947	6	0.6	1849.11	P70206	HEPATIC RECEPTOR HOMOL	5
948	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
949	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
950	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
951	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
952	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
953	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
954	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
955	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
956	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
957	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
958	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
959	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
960	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
961	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
962	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
963	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
964	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
965	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
966	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5
967	6	0.6	1859.11	P70206	HEPATIC RECEPTOR HOMOL	5

[illegible][illegible]

Match	9: Conservative	0: Mismatches	0: Indels	0: Gaps
DB	8 Largest 16			
Cv	119 Largest 117			
RESULT	11			
ID	064730	PRELIMINARY:	PRT: 1343 AA.	
AC	064730:			
DT	01-JUN-1998 (TIMESTAMP, 01, CREATED)			
DP	01-JUN-1998 (TIMESTAMP, 01, LAST SEQUENCE UPDATE)			
DR	01-JUN-1998 (TIMESTAMP, 06, LAST ANNOTATION UPDATE)			
DE	LABECOCCYE COMMON ANTIGEN (L-C) (PDBACCTM):			
OG	MOS MUSCULUS (MOUSE)			
OC	BRANCOFON, ANTONIO, GORDIANO, VERTEDARON, TETRAPPOON, MANAPALIA:			
NR	[1]			
NR	BRANCOFON, ANTONIO.			
RP	SEQUENCE OF 1-1335 FROM N.A.			
RP	MODIFIER: 87203686			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 84:5360-5363(1987).			
NR	[2]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			
RP	REVISION: 1			
RP	PROC. NAT'L ACADE. SCI. U.S.A. 86:6220-6223(1989).			
NR	[1]			
NR	PARENTAL SEQUENCE FROM N.A.			
RP	MODIFIER: 85197328			

[illegible]

PC SEQUENCE FROM N.A.
 RP STRAIN-B*71V.
 RA MEDLINE: 90219504.
 RX MEDLINE: 90219504.
 RL J. VIROL. 64:2064-2072(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 9114498.
 RX MEDLINE: 9114498.
 RL VIROLOGY 181:251-257(1992).
 RN [6]
 RP SEQUENCE FROM N.A.
 RA STRAIN-B*71V.
 RX STRAIN-B*71V.
 RL J. GEN. VIROL. 0:0-0(0).
 RN [7]
 RA ALMAZAN P., RODRIGUEZ J.R., RODRIGUEZ J.M., LA VEGA I., VINUELA E.;
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 91407110.
 RX STRAIN-B*71V.
 RL J. VIROL. 68:2746-2751(1994).
 RN [8]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 93146971.
 RX MEDLINE: 93146971.
 RL YANZ R. J., RODRIGUEZ J. M., RODRIGUEZ J. F., SALAS M. L., VINUELA E.;
 RN J. GEN. VIROL. 74:1653-1658(1993).
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 94065656.
 RX STRAIN-B*71V.
 RL [9]
 RA ALCAÑI M., ANGULO A.; VINUELA E.;
 RP J. GEN. VIROL. 74:2317-2324(1993).
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 93227388.
 RX MEDLINE: 93227388.
 RL M. SALAS M. L., VINUELA E., LOPEZ-OTIN C.;
 RN ARCH. VIROL. 130:93-107(1993).
 RP [11]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 9337780.
 RX MEDLINE: 9337780.
 RL BIACRO R., LOPEZ-OTIN C., MONOZ M., BOCKAMP E. O., SINOR-VARDA C.;
 RN VINUELA E.;
 RL J. GEN. VIROL. 178:301-304(1990).
 RP [12]
 RP SEQUENCE FROM N.A.
 RA STRAIN-B*71V.

RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.M., NOGAL M.L., YOSTE L., VINUELA E.;
 R NUCLEIC ACIDS RES. 21:443-447(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.M., ALAMAN F., VINUELA E., RODRIGUEZ J.F.;
 R J. VIROL. 67:511-520(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.M., BOURNELL M., RODRIGUEZ J.F., VINUELA E.;
 R GENE 134:161-174(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9023393.
 RA LOPES-OTIN C., FREIZE J.M., PARRA F., MENDEZ E., VINUELA E.;
 R VIROLOGY 175:477-484(1990).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9413396.
 RA YANEZ R.J., YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;
 R GENE 158:103-110(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;
 R J. VIROL. 67:2475-2485(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313310.
 RA PRADOS P.J., VINUELA E., ALCAMI A.;
 R J. VIROL. 67:2475-2485(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9226060.
 RA ALCAMI A., ANGULO A., LOPES-OTIN C., MONES M., FREIZE J.M.;
 R J. VIROL. 66:3860-3868(1992).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., VINUELA E.;
 R NUCLEIC ACIDS RES. 21:443-447(1993).
 RP SEQUENCE FROM N.A.

SQ SEQUENCE 146 AA: 17258 MW: 83574504 CRC32:
 Query Match 0.84: Score 8: DB 14: Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;
 DB 7 RESIDUARY 14
 QY 142 RESIDUARY 149
 RESULT 13 PRELIMINARY: PRT: 338 AA.
 AC 065547;
 DT 01-NOV-1996 (TREMBL:REL. 01, CREATED)
 DT 01-NOV-1996 (TREMBL:REL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL:REL. 07, LAST ANNOTATION UPDATE)
 D0 01-NOV-1996
 OS BOVINE HERPESVIRUS TYPE 1.
 OC VIRIDAE: DS-DNA ENVELOPED VIRUSES: HERPESVIRINAE: ALPHAHERPESVIRINAE.
 RN SEQUENCE FROM N.A.
 RA KRODOR A., TIRIO S.F., BABIUX L.A., WOKS S.;
 R L. GENE 0.0:0.0(0).
 SQ SEQUENCE 338 AA: 35930 MW: 47818424 CRC32:
 Query Match 0.84: Score 8: DB 14: Length 338;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;
 DB 270 ARRRGSA 277
 QY 243 ARRRGSA 250
 RESULT 14 PRELIMINARY: PRT: 342 AA.
 AC 083745;
 DT 01-NOV-1996 (TREMBL:REL. 01, CREATED)
 DT 01-NOV-1996 (TREMBL:REL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL:REL. 06, LAST ANNOTATION UPDATE)
 D0 01-NOV-1996
 OS PROVIRAL V-MOS.
 OC METEOROLATPATIVE SARCOMA VIRUS.
 OS VIRIDAE: DS-DNA ENVELOPED VIRUSES: POSITIVE-STRAND, RETROVIRINAE.
 RN SEQUENCE FROM N.A.
 RA STINER A., ARNDTJOTT C., KOLLER R., COGGINS L., OSTERING N.;
 R J. VIROL. 67:511-520(1993).
 RP SEQUENCE FROM N.A.

RL VIROLOGY 193:531-536(1993).
 RN SEQUENCE FROM N.A.
 RX MEDLINE: 93174941.
 RA PERA L., YANEZ R.J., REVILLA Y., VINUELA E., SALAS M.L.;
 R VIROLOGY 193:531-536(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 93159428.
 RA YANEZ R.J., FREIZE J.M., MONES G., LOPES-OTIN C., VINUELA E.;
 R J. VIROL. 67:511-520(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;
 R GENE 134:161-174(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9023393.
 RA LOPES-OTIN C., FREIZE J.M., PARRA F., MENDEZ E., VINUELA E.;
 R VIROLOGY 175:477-484(1990).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9413396.
 RA YANEZ R.J., YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;
 R GENE 158:103-110(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., RODRIGUEZ J.F., VINUELA E., SALAS M.L.;
 R J. VIROL. 67:2475-2485(1993).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9226060.
 RA ALCAMI A., ANGULO A., LOPES-OTIN C., MONES M., FREIZE J.M.;
 R J. VIROL. 66:3860-3868(1992).
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 9313190.
 RA YANEZ R.J., VINUELA E.;
 R NUCLEIC ACIDS RES. 21:443-447(1993).
 RP SEQUENCE FROM N.A.

RL J. VIROL. 50:725-732(1984).
 DR PROSTATE: 5001081 PROTEIN: KINASE, DT: 1.
 DR PROSTATE: 5001081 PROTEIN: KINASE, DT: 1.
 RM COAT PROTEIN: ENVELOPE PROTEIN.
 SQ SEQUENCE 342 AA: 37871 MW: 90793921 CRC32:
 Query Match 0.84: Score 8: DB 14: Length 342;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;
 DB 291 LABSLON 298
 QY 372 LABSLON 379
 RESULT 15 PRELIMINARY: PRT: 508 AA.
 AC 013354;
 DT 01-NOV-1997 (TREMBL:REL. 04, CREATED)
 DT 01-NOV-1997 (TREMBL:REL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (TREMBL:REL. 06, LAST ANNOTATION UPDATE)
 D0 01-NOV-1997
 OS GALINUS GALUS (CHICKEN).
 OC VIRIDAE: DS-DNA ENVELOPED VIRUSES: HERPESVIRINAE: ALPHAHERPESVIRINAE.
 RN SEQUENCE FROM N.A.
 RA KRODOR A., TIRIO S.F., BABIUX L.A., WOKS S.;
 R L. GENE 0.0:0.0(0).
 SQ SEQUENCE 508 AA: 58195 MW: 88396404 CRC32:
 Query Match 0.84: Score 8: DB 13: Length 508;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0;
 DB 128 PROVIRAL 135
 QY 436 PROVIRAL 433
 RESULT 16 PRELIMINARY: PRT: 680 AA.
 AC 015740;
 DT 01-NOV-1997 (TREMBL:REL. 01, CREATED)
 DT 01-NOV-1997 (TREMBL:REL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (TREMBL:REL. 06, LAST ANNOTATION UPDATE)
 D0 01-NOV-1997
 OS PROVIRAL V-MOS.
 OC METEOROLATPATIVE SARCOMA VIRUS.
 OS VIRIDAE: DS-DNA ENVELOPED VIRUSES: POSITIVE-STRAND, RETROVIRINAE.
 RN SEQUENCE FROM N.A.
 RA STINER A., ARNDTJOTT C., KOLLER R., COGGINS L., OSTERING N.;
 R J. VIROL. 67:511-520(1993).
 RP SEQUENCE FROM N.A.

AC 015740: (PRELIMINARY: 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-JAN-1998 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE PKSA (FRAGMENT).
 GN PKSA (FRAGMENT).
 OS DROSOPHILA DISCOIDEUM (FRUIT FLY).
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 680 AA: 76448 MW; 2828560 CMC32;
 Query Match 0.84; Score 8; DB 5; Length 680;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 166 REELAF 173
 QY 359 REELAF 566
 RESULT 17
 ID 024144: (PRELIMINARY: PRT: 869 AA.
 AC 024144: (PRELIMINARY: 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE INTRONED PROTEIN.
 GN INTRONED PROTEIN.
 OS DROSOPHILA MELANOGASTER (FRUIT FLY).
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 869 AA: 97869 MW; 2395888 CMC32;
 Query Match 0.84; Score 8; DB 5; Length 869;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 430 SRSPLPX 437

RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 23 AA: 2706 MW; A872074 CMC32;
 Query Match 0.74; Score 7; DB 5; Length 23;
 Best Local Similarity 100.0%; Pred. No. 1.01e-01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 7 VERTUL 13
 QY 785 VERTUL 791
 RESULT 20
 ID 065832: (PRELIMINARY: PRT: 37 AA.
 AC 065832: (PRELIMINARY: 07, CREATED)
 DT 01-NOV-1996 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE Lycopodium 4.2 NO PROTEIN (FRAGMENT).
 GN Lycopodium 4.2 NO PROTEIN (FRAGMENT).
 OS LYCOPODIUM ESCULENTUM (TOMATO).
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 37 AA: 4405 MW; 5762180 CMC32;
 Query Match 0.74; Score 7; DB 10; Length 37;
 Best Local Similarity 100.0%; Pred. No. 1.01e-01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 5 RSPLFS 11
 QY 364 RSPLFS 370
 RESULT 21

QY 251 SRSPLPX 258
 RESULT 18
 ID 060300: (PRELIMINARY: PRT: 1095 AA.
 AC 060300: (PRELIMINARY: 07, CREATED)
 DT 01-NOV-1996 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE KIN5053 PROTEIN (FRAGMENT).
 GN KIN5053 PROTEIN (FRAGMENT).
 OS KIN5053 PROTEIN (FRAGMENT).
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 1095 AA: 118835 MW; 4347424 CMC32;
 Query Match 0.84; Score 8; DB 4; Length 1095;
 Best Local Similarity 100.0%; Pred. No. 1.03e-01;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 941 ALAPTRA 948
 QY 15 ALAPTRA 22
 RESULT 19
 ID 027174: (PRELIMINARY: PRT: 23 AA.
 AC 027174: (PRELIMINARY: 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE TRICHOCEST MATRIX PROTEIN T3 (FRAGMENT).
 GN TRICHOCEST MATRIX PROTEIN T3 (FRAGMENT).
 OS PARAMECELETTA TETRAURELIA.
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 23 AA: 2706 MW; A872074 CMC32;
 Query Match 0.74; Score 7; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.01e-01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 52 LVAPSCA 58
 QY 183 LVAPSCA 189

RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 87 AA: 9613 MW; 3825180 CMC32;
 Query Match 0.74; Score 7; DB 2; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.01e-01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 52 LVAPSCA 58
 QY 183 LVAPSCA 189
 RESULT 22
 ID 031890: (PRELIMINARY: PRT: 109 AA.
 AC 031890: (PRELIMINARY: 01, CREATED)
 DT 01-NOV-1996 (TREMBLREL: 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBLREL: 05, LAST ANNOTATION UPDATE)
 DE BMTA (FRAGMENT).
 GN BMTA (FRAGMENT).
 OS BMTA (FRAGMENT).
 OC EUTAROTIA: PROTODIA: SARCOMASTICOPHORA: SARCODINIA: RHIZOPODA:
 RN (1)
 RP SEQUENCE FROM N.A.
 RA LOMIS M.F., JAMES N.;
 RL SUBMITTED (AUG-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF019983; G435115; -
 FT NON_TER 1
 SQ SEQUENCE 109 AA: 11728 MW; 4347424 CMC32;
 Query Match 0.74; Score 7; DB 3; Length 109;
 Best Local Similarity 100.0%; Pred. No. 1.01e-01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 5 RSPLFS 11
 QY 364 RSPLFS 370
 RESULT 21

DB 63 GASTRUL 69
 699 GASTRUL 705
 640 TSELP 646

RESULT 23 PRELIMINARY: PRT: 111 M.
 ID Q04293
 AC Q04293
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE HYPOTHETICAL 12.3 ED PROTEIN
 GN ORF 2
 OS HANSENULA POLYOMORPHA (TEAR)
 OC EUBACTERIA: FORCI: ACOETICOTIN, HERIACONCECTES.
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 111 M.; 12291 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 22 TSELP 28
 640 TSELP 646

RESULT 24 PRELIMINARY: PRT: 113 M.
 ID Q03100
 AC Q03100
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE ORF27 (FRAGMENT)
 OS RHODOSPIRILLACEAE: RHODOSPIRILLACEAE
 OC RHODOSPIRILLACEAE: ANOXYPHOTOBACTERIA: PURPLE BACTERIA:
 RP SEQUENCE FROM N.A.
 RC STRAIN-2.4.1;
 RA ZILINSKIS, J. H.; KAPLAN S.;
 DR EMBL: U15463 (C1742) (1993)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 117 M.; 12770 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OC VIRIDAE: SS-RNA ENVELOPED VIRUSES: POSITIVE-STRAND: RETROVIRIDAE;
 OC LENTIVIRIDAE
 RP SEQUENCE FROM N.A.
 RC STRAIN-N;
 RA XIANG Z., AITOCHI K., WILKINS A., DIAS P., WHITTLE H., BREWER J.;
 DR EMBL: A100812 (C17354)
 DE HYPOTHETICAL 14.4 NO PROTEIN
 GN ORF 2
 OS HANSENULA POLYOMORPHA (TEAR)
 OC EUBACTERIA: FORCI: ACOETICOTIN, HERIACONCECTES.
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 111 M.; 12291 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 7 PSTRP 13
 330 PSTRP 336

RESULT 27 PRELIMINARY: PRT: 128 M.
 ID Q05981
 AC Q05981
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE HYPOTHETICAL 14.4 NO PROTEIN
 GN ORF 2
 OS HANSENULA POLYOMORPHA (TEAR)
 OC EUBACTERIA: FORCI: ACOETICOTIN, HERIACONCECTES.
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 111 M.; 12291 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 46 AVIRID 52
 811 AVIRID 817

RESULT 28 PRELIMINARY: PRT: 135 M.
 ID Q08029
 AC Q08029
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOMA-LYMPHOMA VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT)
 GN TAX-2
 OS HUMAN T-CELL LYMPHOTROPIC VIRUS TYPE II
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP;
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 117 M.; 12770 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 65 ABRGCS 71
 243 ABRGCS 249

RESULT 29 PRELIMINARY: PRT: 138 M.
 ID Q08032
 AC Q08032
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOMA-LYMPHOMA VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT)
 GN TAX-2
 OS HUMAN T-CELL LYMPHOTROPIC VIRUS TYPE II
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP;
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 117 M.; 12770 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 138;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 46 AVIRID 52
 811 AVIRID 817

RESULT 28 PRELIMINARY: PRT: 135 M.
 ID Q08029
 AC Q08029
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOMA-LYMPHOMA VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT)
 GN TAX-2
 OS HUMAN T-CELL LYMPHOTROPIC VIRUS TYPE II
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP;
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 117 M.; 12770 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 72 SHELUP 78
 251 SHELUP 257

RESULT 29 PRELIMINARY: PRT: 138 M.
 ID Q08032
 AC Q08032
 DT 01-NOV-1996 (TREMBL, 01, CREATED)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TREMBL, 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOMA-LYMPHOMA VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT)
 GN TAX-2
 OS HUMAN T-CELL LYMPHOTROPIC VIRUS TYPE II
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP;
 RP SEQUENCE FROM N.A.
 RC STRAIN-DL1
 RA KUTYLINA A.I., SERGINIA S.A., TIKHONOVA L.P., KRYUKOV V.N.;
 DR EMBL: X5863 (C1744)
 DE HYPOTHETICAL PROTEIN
 SQ SEQUENCE 117 M.; 12770 M.; 394881A CIRC32;
 Query Match 0.78; Score 7; DB 3; Length 138;
 Best Local Similarity 100.0%; Pred. No. 1.0e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible][illegible]

```

EX MEDLINE: 95286574.
RA HODGKINSON A. A., ROSOVIST R., WOLF-WANTZ H., FORBESNG A.,
RA REEVE J. C. LAMON. 63:289-323(6/1995).
RN [2]
RP SEQUENCE FROM N. A.
RA HODGKINSON A. A., 1994) NO DBS//GENBANK/DBS DATA BANKS.
DB EMBL: D18804, C619941.
SV PLASMID.
SQ SEQUENCE 182 AA: 20959 MW: C74A6FD CRC32:
Query Match 0.7% Score 7; DB 2; Length 182;
Best Local Similarity 100.0%; Pcd. No. 1.01e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
Db 63 GSI:STL 68
OY 856 GSI:STL 862

RESULT 44 PRELIMINARY: PRT: 182 AA.
ID 068704.
AC 068704.188 (TRIMBDEL. 07 CREATED)
DE 01-NOV-1999 (TRIMBDEL. 07) LAST SEQUENCE UPDATE)
D7 01-NOV-1999 (TRIMBDEL. 07, LAST ANNOTATION UPDATE)
D8 TOP (TOP) HOMOLOG (TOP).
DE YREINIA PESTIS.
OS
OC PLASMID PCOI.
CC PROKARYOT. CIRCLICUTES; SCOTOBACTERIA; FACULTATIVELY ANAEROBIC RODS
CC ENTEROBACTERIALE.
RN [1]
RP SEQUENCE FROM N. A.
RA STRAIN-KIM.
RA HU P., ELLIOTT J., MCKENDRY P., SEMONSKI E., GARNES J.,
RL SCHMITZ (MAR-1998) NO DBS//GENBANK/DBS DATA BANKS.
DB EMBL: F050596; G2895260.
SV PLASMID.
SQ SEQUENCE 182 AA: 21000 MW: ADH41D7 CRC32:
Query Match 0.7% Score 7; DB 2; Length 182;
Best Local Similarity 100.0%; Pcd. No. 1.01e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
Db 63 GSI:STL 68
OY 856 GSI:STL 862

```

[illegible]

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

WIPEROLIN

(TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

March: protein - protein database search, using Smith-Waterman algorithm

Run on: F71 Dec 18 19:04:41 1998; Nmap time 37.18 Seconds

Tabular output not generated.

Title: >US-08-951-733-20

Description: (1-1154) from OS08951733.pep

Sequence: 1 HNSGRCVLRHVALPVR.....TLKMAAPLPSDFRITLD 1154

Scoring table: TABLE uniprottable

Gap 60

Searched: 11922 seqs, 1610660 residues

Post-processing: Minimum Match 0%

Listing first 1000 summaries

Database: 4-geneeq2
1:part1 2:part3 3:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29

Statistics: Mean 3.325; Variance 0.726; scale 4.592

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

48	6	0.5	17 25	W17555	Beta-B2-crystallin fr	5.54e+02
49	6	0.5	20 28	W41112	Tax protein fragment	5.54e+02
50	6	0.5	20 14	R14126	HIV-1 p04(tax) prote	5.54e+02
51	6	0.5	20 14	R14126	HIV-1 p04(tax) prote	5.54e+02
52	6	0.5	21 13	R73929	B-coetaxial CD epit	5.54e+02
53	6	0.5	23 27	W35875	Leader sequence for u	5.54e+02
54	6	0.5	23 29	R52639	Myotrophin typtic p	5.54e+02
55	6	0.5	23 29	R52639	Myotrophin typtic p	5.54e+02
56	6	0.5	23 29	R52639	Myotrophin typtic p	5.54e+02
57	6	0.5	26 19	W03637	C-protein-coupled odo	5.54e+02
58	6	0.5	28 21	W11332	Glucagon like peptide	5.54e+02
59	6	0.5	31 20	W03890	RNA region of Rcy 22	5.54e+02
60	6	0.5	33 23	W04489	Peptide 2 of the alph	5.54e+02
61	6	0.5	33 28	W4502	Human glucokinase exo	5.54e+02
62	6	0.5	33 28	W4502	Human glucokinase exo	5.54e+02
63	6	0.5	45 22	W20450	Human glucokinase exo	5.54e+02
64	6	0.5	53 8	R42167	Human glucokinase exo	5.54e+02
65	6	0.5	53 8	R42166	Human glucokinase exo	5.54e+02
66	6	0.5	54 15	R13599	ORF encoded by exon 6	5.54e+02
67	6	0.5	54 15	R13599	ORF encoded by exon 6	5.54e+02
68	6	0.5	54 15	R13599	ORF encoded by exon 6	5.54e+02
69	6	0.5	62 22	W00316	R. ptylori secreted or	5.54e+02
70	6	0.5	62 22	W00316	R. ptylori secreted or	5.54e+02
71	6	0.5	66 20	W10172	R. ptylori ORF 14e103	5.54e+02
72	6	0.5	68 29	W55293	Adeno acid 284-351 o	5.54e+02
73	6	0.5	70 16	R10806	B-coetaxial CD epit	5.54e+02
74	6	0.5	71 13	R12658	Human immunoglobulin	5.54e+02
75	6	0.5	75 24	W26420	Sytlepox Virus HindII	5.54e+02
76	6	0.5	76 29	W44765	Sequence encoded by t	5.54e+02
77	6	0.5	88 23	W12450	Myobacterium tubercu	5.54e+02
78	6	0.5	88 23	W12450	Myobacterium tubercu	5.54e+02
79	6	0.5	88 23	W12450	Myobacterium tubercu	5.54e+02
80	6	0.5	94 10	R50933	Sequence of the first	5.54e+02
81	6	0.5	95 2	P80411	Sequence of the first	5.54e+02
82	6	0.5	95 2	P80411	Sequence of the first	5.54e+02
83	6	0.5	103 23	W24695	R. ptylori secreted or	5.54e+02
84	6	0.5	103 23	W24695	R. ptylori secreted or	5.54e+02
85	6	0.5	103 23	W24695	R. ptylori secreted or	5.54e+02
86	6	0.5	109 13	P60796	Fragment of sak gene	5.54e+02
87	6	0.5	113 12	W04547	Human immunoglobulin	5.54e+02
88	6	0.5	113 12	W04547	Human immunoglobulin	5.54e+02
89	6	0.5	118 12	R63440	Human immunoglobulin	5.54e+02
90	6	0.5	118 12	R63440	Human immunoglobulin	5.54e+02
91	6	0.5	118 12	R63440	Human immunoglobulin	5.54e+02
92	6	0.5	118 12	R63440	Human immunoglobulin	5.54e+02
93	6	0.5	118 12	R63440	Human immunoglobulin	5.54e+02
94	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
95	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
96	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
97	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
98	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
99	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02
100	6	0.5	123 23	W08505	Heavy chain 42 for an	5.54e+02

Result	Score	Query	Match	Length	ID	Description	Pred. No
No.							
1	7	0.6	126 9	R46892	CcBb cytotoxin fused	4.52e+01	
2	7	0.6	136 20	W03091	Staphylokinase deriva	4.52e+01	
3	7	0.6	146 12	R41895	Sequence of human gta	4.52e+01	
4	7	0.6	146 12	R41895	Rex region gene produ	4.52e+01	
5	7	0.6	296 25	W05397	Human clone 55 protei	4.52e+01	
6	7	0.6	333 23	W19736	Sugar biosynthesis en	4.52e+01	
7	7	0.6	334 1	R45741	Myosinotol dehydroge	4.52e+01	
8	7	0.6	334 1	R45741	Myosinotol dehydroge	4.52e+01	
9	7	0.6	355 11	R00950	Recombinant high effi	4.52e+01	
10	7	0.6	355 6	R28972	Sequence in a high at	4.52e+01	
11	7	0.6	358 29	W44266	Human HPA proteoglyca	4.52e+01	
12	7	0.6	358 20	W08135	Human cytokine respon	4.52e+01	
13	7	0.6	395 14	R00288	galk gene of S. livid	4.52e+01	
14	7	0.6	395 14	R00288	galk gene of S. livid	4.52e+01	
15	7	0.6	395 14	R00288	galk gene of S. livid	4.52e+01	
16	7	0.6	395 14	R00288	Sequence encoded by S	4.52e+01	
17	7	0.6	398 26	W20252	Ammonium ferritin	4.52e+01	
18	7	0.6	439 17	R08407	beta-1.65-endoglucan	4.52e+01	
19	7	0.6	443 13	R08407	beta-1.65-endoglucan	4.52e+01	
20	7	0.6	493 13	R72470	JobbA fatty acyl-CoA	4.52e+01	
21	7	0.6	493 18	R87508	JobbA fatty acyl-red	4.52e+01	
22	7	0.6	493 5	R26898	JobbA fatty acyl-red	4.52e+01	
23	7	0.6	493 14	R71954	JobbA fatty acyl-red	4.52e+01	
24	7	0.6	493 14	R71954	Respirated JobbA	4.52e+01	
25	7	0.6	493 14	R87509	Respirated JobbA	4.52e+01	
26	7	0.6	517 27	W03099	Tobacco calcium/calmo	4.52e+01	
27	7	0.6	834 1	R04869	Alpha-1.6-glucanase	4.52e+01	
28	7	0.6	894 2	R72127	Phosphatidyl-coester	4.52e+01	
29	7	0.6	919 1	R05512	Phosphatidyl-coester	4.52e+01	
30	7	0.6	919 1	R05512	Phosphatidyl-coester	4.52e+01	
31	7	0.6	946 26	W27085	Human cytochromeP450	4.52e+01	
32	7	0.6	946 36	W27085	Human cytochromeP450	4.52e+01	
33	7	0.6	986 14	R70505	Heat shock protein	4.52e+01	
34	7	0.6	1062 8	R19055	Heat shock protein	4.52e+01	
35	7	0.6	1396 21	W19055	Human SHC protein	4.52e+01	
36	7	0.6	1577 16	R19047	Alpha-D-glucosyltran	4.52e+01	
37	6	0.5	5 15	R83542	EL2684 zinc finger 2	5.52e+02	
38	6	0.5	5 15	R83542	EL2684 zinc finger 2	5.52e+02	
39	6	0.5	10 32	W05577	Peptide protein kinase	5.52e+02	
40	6	0.5	10 32	W05577	T-cell modulating pep	5.52e+02	
41	6	0.5	12 33	R18186	Peptide betaactyl for	5.52e+02	
42	6	0.5	13 35	R19346	GZF-SRC protein tyros	5.52e+02	
43	6	0.5	13 37	W42500	Biologically active	5.52e+02	
44	6	0.5	15 37	W42500	Biologically active	5.52e+02	
45	6	0.5	15 37	W42500	Biologically active	5.52e+02	
46	6	0.5	15 36	R63100	Peptide recombining	5.52e+02	
47	6	0.5	15 36	R63100	Peptide recombining	5.52e+02	

50 Sequence 375 AA:
Query Match
Best Local Similarity 100.0%: Pred. No. 4,52e+01;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 260 EPIPAD 266
|||||
QY 804 EPIPAD 810

RESULT 14
ID R80288 standard: Protein: 395 AA.
AC R80288;
DT 21-DEC-1995 (first entry)
DE 21-DEC-1995 (first entry)
NM Galactose metabolism, enzyme, galK, gal operon; galactokinase.
OS Streptomyces lividans.
PN 055435730-A.
PR 25-DEC-1993: 834706.
PR 28-FEB-1986: US-934706.
PR 30-JAN-1987: US-009419.
PR 31-APR-1991: US-692769.
PR 15-DEC-1992: US-692769.
PR 30-DEC-1993: US-009419.
PA (ADAM) ADAMS C W.
PA (BRAY) BRAY W E.
PA (FOUNT) FOUNT J A.
PA (KAM) KAMMER R J.
PA (KAM) KAMMER R J, ME, Formwald JN, Schmidt FJ;
DR WP1: 95-26848/75.
DR N-PDB: Q98377.
PT Recombinant DNA molecules for producing expression/cloning vectors in
Example: Table 1: 2pp; English.
CC A recombinant DNA molecule comprising a Streptomyces spp. gal operon
CC promoter or any regulatable and functional deletion derivative.
CC promoter, galT, galZ and galK structural genes and any other
CC regulatory regions required for their transcription and translation.
CC galT is uridine diphosphoglucose-4-epimerase; galZ is
CC galactokinase; and galK is galactose-1-phosphate uridylyltransferase.
CC The gal operon is located on a 1.1 kb DNA fragment, which is a
CC region of Streptomyces lividans 1126. It was sequenced by chain
CC termination and its complete DNA sequence is shown in Q98377.
CC The p1 promoter is galactose inducible, glucose repressible and is
CC the regulatory promoter for the gal operon. Data indicates
CC that the p1 promoter is a hybrid of the gal operon and the
CC kb. The transcript consists of approx. 1kb each for galT, galZ and

50 Sequence 395 AA:
Query Match
Best Local Similarity 100.0%: Pred. No. 4,52e+01;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 375 PIRGAP 381
|||||
QY 260 PIRGAP 266

RESULT 15
ID P70276 standard: Protein: 395 AA.
AC P70276;
DT 28-FEB-1992 (first entry)
DE Sequence encoded by S. lividans gal operon galK gene.
NM Galactose metabolism, promoter; expression vector; gal operon.
OS Streptomyces lividans.
PN 055435730-A. lividans strain 1126.
PR 07-SEP-1987: 870924.
PR 28-FEB-1986: US-934706.
PR 30-JAN-1987: US-009419.
PA (SMK) SHINKLINE BECKMAN CORP.
PA (ADAM) ADAMS C W, Formwald JN, Schmidt FJ;
DR WP1: 95-26848/75.
PT Recombinant DNA contg. Streptomyces gal operon - or its
PT promoter and gene fragments, and transformed cell able to express
PT foreign protein and to metabolize galactose
PT (recombinant; Page 33,11, P70276). English
CC The gal operon is located on a 1.1 kb DNA fragment, which is a
CC region of Streptomyces lividans 1126. It was sequenced by chain
CC termination and its complete DNA sequence is shown in Q98377.
CC The p1 promoter is galactose inducible, glucose repressible and is
CC the regulatory promoter for the gal operon. Data indicates
CC that the p1 promoter is a hybrid of the gal operon and the
CC kb. The transcript consists of approx. 1kb each for galT, galZ and

CC galK. Galactose induction of p1 is mediated, at least in part, by
CC transcription start site and a repressor protein which recognizes
CC the operator. The p2 promoter is upstream of the galT gene and
CC transcribes both galT and galK genes. p2 promoter expression is
CC constitutive.
50 Sequence 395 AA:
Query Match
Best Local Similarity 100.0%: Pred. No. 4,52e+01;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 375 PIRGAP 381
|||||
QY 260 PIRGAP 266

RESULT 15
ID R41530 standard: Protein: 395 AA.
AC R41530;
DT 28-FEB-1996 (first entry)
DE 28-FEB-1996 (first entry)
NM Galactose metabolism, enzyme, galK, gal operon; galactokinase.
OS Streptomyces lividans.
PN 055435730-A.
PR 25-DEC-1993: 834706.
PR 28-FEB-1986: US-934706.
PR 30-JAN-1987: US-009419.
PR 31-APR-1991: US-692769.
PR 15-DEC-1992: US-692769.
PR 30-DEC-1993: US-009419.
PA (ADAM) ADAMS C W.
PA (BRAY) BRAY W E.
PA (FOUNT) FOUNT J A.
PA (KAM) KAMMER R J.
PA (KAM) KAMMER R J, ME, Formwald JN, Schmidt FJ;
DR WP1: 95-26848/75.
DR N-PDB: Q98377.
PT Recombinant DNA molecules for producing expression/cloning vectors in
Example: Table 1: 2pp; English.
CC A recombinant DNA molecule comprising a Streptomyces spp. gal operon
CC promoter or any regulatable and functional deletion derivative.
CC promoter, galT, galZ and galK structural genes and any other
CC regulatory regions required for their transcription and translation.
CC galT is uridine diphosphoglucose-4-epimerase; galZ is
CC galactokinase; and galK is galactose-1-phosphate uridylyltransferase.
CC The gal operon is located on a 1.1 kb DNA fragment, which is a
CC region of Streptomyces lividans 1126. It was sequenced by chain
CC termination and its complete DNA sequence is shown in Q98377.
CC The p1 promoter is galactose inducible, glucose repressible and is
CC the regulatory promoter for the gal operon. Data indicates
CC that the p1 promoter is a hybrid of the gal operon and the
CC kb. The transcript consists of approx. 1kb each for galT, galZ and

50 Sequence 395 AA:
Query Match
Best Local Similarity 100.0%: Pred. No. 4,52e+01;
Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 157 IIRAYP 163
|||||
QY 437 IIRAYP 443

RESULT 18
ID R88407 standard: Protein: 439 AA.
AC R88407;
DT 05-DEC-1996 (first entry)
DE 05-DEC-1996 (first entry)
NM Galactose metabolism, enzyme, galK, gal operon; galactokinase.
OS Streptomyces lividans.
PN 055435730-A.
PR 25-DEC-1993: 834706.
PR 28-FEB-1986: US-934706.
PR 30-JAN-1987: US-009419.
PR 31-APR-1991: US-692769.
PR 15-DEC-1992: US-692769.
PR 30-DEC-1993: US-009419.
PA (ADAM) ADAMS C W.
PA (BRAY) BRAY W E.
PA (FOUNT) FOUNT J A.
PA (KAM) KAMMER R J.
PA (KAM) KAMMER R J, ME, Formwald JN, Schmidt FJ;
DR WP1: 95-26848/75.
DR N-PDB: Q98377.
PT Recombinant DNA molecules for producing expression/cloning vectors in
Example: Table 1: 2pp; English.
CC A recombinant DNA molecule comprising a Streptomyces spp. gal operon
CC promoter or any regulatable and functional deletion derivative.
CC promoter, galT, galZ and galK structural genes and any other
CC regulatory regions required for their transcription and translation.
CC galT is uridine diphosphoglucose-4-epimerase; galZ is
CC galactokinase; and galK is galactose-1-phosphate uridylyltransferase.
CC The gal operon is located on a 1.1 kb DNA fragment, which is a
CC region of Streptomyces lividans 1126. It was sequenced by chain
CC termination and its complete DNA sequence is shown in Q98377.
CC The p1 promoter is galactose inducible, glucose repressible and is
CC the regulatory promoter for the gal operon. Data indicates
CC that the p1 promoter is a hybrid of the gal operon and the
CC kb. The transcript consists of approx. 1kb each for galT, galZ and


```

971 6 0.5 2142 2 B35088 HMC class III histone 5.47e+02
972 6 0.5 2119 1 GNVY44 genome polypeptide - 5.47e+02
973 6 0.5 2119 1 GNVY44 genome polypeptide - 5.47e+02
974 6 0.5 2225 1 A23443 pyridoxal synthase 5.47e+02
975 6 0.5 2227 1 GNVY44 genome polypeptide - 5.47e+02
976 6 0.5 2227 1 GNVY44 genome polypeptide - 5.47e+02
977 6 0.5 2227 1 GNVY44 genome polypeptide - 5.47e+02
978 6 0.5 2227 1 GNVY44 genome polypeptide - 5.47e+02
979 6 0.5 2233 1 A39086 calcium channel Bi-1 5.47e+02
980 6 0.5 2316 2 A39086 calcium channel Bi-1 5.47e+02
981 6 0.5 2340 2 A39086 calcium channel Bi-1 5.47e+02
982 6 0.5 2428 2 A39086 calcium channel Bi-1 5.47e+02
983 6 0.5 2428 2 A39086 calcium channel Bi-1 5.47e+02
984 6 0.5 2428 2 A39086 calcium channel Bi-1 5.47e+02
985 6 0.5 2440 2 A39086 calcium channel Bi-1 5.47e+02
986 6 0.5 2440 2 A39086 calcium channel Bi-1 5.47e+02
987 6 0.5 2440 2 A39086 calcium channel Bi-1 5.47e+02
988 6 0.5 2514 1 GNVY44 genome polypeptide - 5.47e+02
989 6 0.5 2567 2 A49551 chrysothrin precursor 5.47e+02
990 6 0.5 2567 2 A49551 chrysothrin precursor 5.47e+02
991 6 0.5 3423 1 A28266 357K segment polypeptide 5.47e+02
992 6 0.5 3423 1 A28266 357K segment polypeptide 5.47e+02
993 6 0.5 3432 1 GNVY44 genome polypeptide - 5.47e+02
994 6 0.5 3432 1 GNVY44 genome polypeptide - 5.47e+02
995 6 0.5 3432 1 GNVY44 genome polypeptide - 5.47e+02
996 6 0.5 3432 1 GNVY44 genome polypeptide - 5.47e+02
997 6 0.5 3898 1 GNVY44 genome polypeptide - 5.47e+02
998 6 0.5 4377 2 A55575 ankyrin 3, long splice 5.47e+02
999 6 0.5 5116 2 A55575 ankyrin 3, long splice 5.47e+02
1000 6 0.5 28926 1 A55575 titin, cardiac muscle 5.47e+02

```

ALIGNMENTS

```

RESULT 1
ENTRY 157644 #type fragment
TITLE transmembrane glycoprotein - mouse (fragment)
FORMAL_NAME Mus musculus (common name house mouse)
DATE 02-Aug-1996 sequence_revision 02-Aug-1996 text_change 02-Aug-1996
ACCESSIONS 157644
REFERENCE 157644
AUTHORS Sage, T.; Tung, J.; Shen, F.W.; Boyce, E.A.
JOURNAL Biol. Cell Biol. (1998) 8:4689-4695
JOURNAL PubMed:9504862
ACROSS-REFERENCES MIM:6090862
ACCESSION 157644
STATUS preliminary; translated from CD/MDL/DDNJ
MOLECULE_TYPE DNA
RESIDUES 1224 #label RES
ACROSS-REFERENCES CB:M23354; MID:9340890; PID:9554192

```

```

RESULT 3
ENTRY A28334 #type complete
TITLE protein-tyrosine-phosphatase (BC 3.1.3.48) Lys-5 precursor
ALTERNATE_NAMES 200K leukocyte common antigen; CD45; PTPN22; T-cell surface
CONTAINS 200K leukocyte common antigen (T-cell variant)
FORMAL_NAME Mus musculus (common name house mouse)
DATE 19-May-1989 sequence_revision 19-May-1989 text_change 19-May-1989
ACCESSIONS A28334; A28381; A60933; A35322; A28075; I54450
REFERENCE A28334
AUTHORS Thomas, M.L.; Reynolds, P.J.; Chalm, A.; Ben-Neriah, Y.;
JOURNAL Trends Biochem. Sci. (1987) 84:5160-5163
JOURNAL PubMed:3046930
TITLE B-cell variant of mouse T200 (Lys-5): evidence for alternative
RNA splicing
ACCESSION A28334
MOLECULE_TYPE RNA
RESIDUES 11291 #label THO
ACROSS-REFERENCES CB:M24455
REFERENCE A28381
AUTHORS Sage, T.; Tung, J.S.; Shen, F.W.; Boyce, E.A.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1986) 83:6940-6944
JOURNAL PubMed:3749314
TITLE Sequences of Lys-5 cDNA: isoform-related diversity of Lys-5
RNA.
ACROSS-REFERENCES MID:8631686
ACCESSION A28381
MOLECULE_TYPE RNA
RESIDUES 1291 #label THO
ACROSS-REFERENCES CB:M15421; MID:9358314; PID:938835
REFERENCE A60933
AUTHORS V.L.; Cleveland, J.L.; Ihle, J.N.
JOURNAL Blood (1991) 78:2212-2218
JOURNAL PubMed:1100000
TITLE Identification of novel protein tyrosine phosphatases of
hematopoietic cells by polymerase chain reaction
amplification
ACROSS-REFERENCES MID:92032882
ACCESSION A61180
MOLECULE_TYPE cDNA
RESIDUES 730-938 #label YIA
ACROSS-REFERENCES CB:M23354
REFERENCE A60933
AUTHORS Gomez, L.I.; Walker, I.D.; Sandlin, M.S.; McKenzie, I.P.C.
JOURNAL J. Biol. Chem. (1991) 266:12463-12468
JOURNAL PubMed:1346381
TITLE High sequence conservation between rat (T200) and mouse

```

```

GENETICS
CLASSIFICATION Lys-5
protein-tyrosine-phosphatase domain homology
KEYWORDS glycoprotein
SIMMARY length 24 checksum 7219
Query Match 0.8; Score 9; DB 2; Length 24;
Best Local Similarity 100.0%; Pred. No. 1,20e-03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 10 LAPCAPLID 18
QY 119 LAPCAPLID 127

```

```

RESULT 2
ENTRY A28335 #type fragment
TITLE leukocyte common antigen precursor - mouse (fragment)
FORMAL_NAME Mus musculus (common name house mouse)
DATE 19-May-1989 sequence_revision 19-May-1989 text_change 19-May-1989
ACCESSIONS A28335
REFERENCE A28335
AUTHORS Sage, T.; Tung, J.S.; Shen, F.W.; Boyce, E.A.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1987) 84:5164-5168
JOURNAL PubMed:3046930
TITLE Alternative use of 5' exon in the specification of Lys-5
protein-tyrosine-phosphatase (BC 3.1.3.48)
ACCESSION A28335
MOLECULE_TYPE RNA
RESIDUES 1183 #label SAG
ACROSS-REFERENCES CB:M24455; MID:9340890; PID:9554192
CLASSIFICATION #superfamily leukocyte common antigen; leukocyte common
protein-tyrosine-phosphatase domain homology
FEATURE 24-183
#domain signal sequence #status predicted label SAG
#product leukocyte common antigen, 200K status
SIMMARY #length 183 checksum 9727
Query Match 0.8; Score 9; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1,20e-03;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 8 LAPCAPLID 16
QY 119 LAPCAPLID 127

```

```

ACCESSION A60933
MOLECULE_TYPE RNA
RESIDUES 730-938 #label YIA
ACROSS-REFERENCES CB:M23354; MID:9340890; PID:9554192
REFERENCE A28075
AUTHORS Ben-Neriah, Y.; Thomas, M.L.; Chalm, A.; Reynolds, P.J.
JOURNAL Trends Biochem. Sci. (1987) 84:5161-5165
JOURNAL PubMed:3046930
TITLE Cloned murine T200 (Lys-5) cDNA reveals multiple transcripts
within B- and T-lymphocyte lineages.
ACROSS-REFERENCES MID:87092355
ACCESSION A35322
MOLECULE_TYPE RNA
RESIDUES 961-1291 #label BAS
ACROSS-REFERENCES CB:M15174; MID:9201105; PID:9201106
REFERENCE I54450
AUTHORS Thomas, M.L.; Reynolds, P.J.; Chalm, A.; Ben-Neriah, Y.
JOURNAL Trends Biochem. Sci. (1987) 84:5160-5163
JOURNAL PubMed:3046930
TITLE B-cell variant of mouse T200 (Lys-5): evidence for alternative
RNA splicing
ACROSS-REFERENCES CB:M24455; MID:9340890; PID:9554192
ACCESSION A28334
MOLECULE_TYPE DNA
RESIDUES 11291 #label THO
ACROSS-REFERENCES CB:M23354; MID:9340890; PID:9554192
REFERENCE A28381
AUTHORS Sage, T.; Tung, J.S.; Shen, F.W.; Boyce, E.A.
JOURNAL Proc. Natl. Acad. Sci. U.S.A. (1986) 83:6940-6944
JOURNAL PubMed:3749314
TITLE Sequences of Lys-5 cDNA: isoform-related diversity of Lys-5
RNA.
ACROSS-REFERENCES MID:8631686
ACCESSION A28381
MOLECULE_TYPE RNA
RESIDUES 1291 #label THO
ACROSS-REFERENCES CB:M15421; MID:9358314; PID:938835
REFERENCE A60933
AUTHORS V.L.; Cleveland, J.L.; Ihle, J.N.
JOURNAL Blood (1991) 78:2212-2218
JOURNAL PubMed:1100000
TITLE Identification of novel protein tyrosine phosphatases of
hematopoietic cells by polymerase chain reaction
amplification
ACROSS-REFERENCES MID:92032882
ACCESSION A61180
MOLECULE_TYPE cDNA
RESIDUES 730-938 #label YIA
ACROSS-REFERENCES CB:M23354
REFERENCE A60933
AUTHORS Gomez, L.I.; Walker, I.D.; Sandlin, M.S.; McKenzie, I.P.C.
JOURNAL J. Biol. Chem. (1991) 266:12463-12468
JOURNAL PubMed:1346381
TITLE High sequence conservation between rat (T200) and mouse

```



```

title      Analysis of a 42.5 kb DNA sequence of chromosome X reveals three RNA genes and a new open reading frame encoding a novel protein belonging to the family of ubiquitin-protein ligases.
 accession #S6789
 molecule_type DMSA, alpha1b1 RNA
 $domain_references EMBL:D0544; NCBI:g139700; PDB:g139703
 $note     The nucleotide sequence was submitted to the EMBL data library, February 1996
 GENETICS
   gene     SCD, SCCL1; ENH1
   $accession_numbers SCD:S0003806; MIPS:XMR045C
 map_position 10R
 function    involved in nuclear
             involvement in protein folding and assembling/disassembling of protein complexes
             specifically heat shock protein 70
             associated with mitochondrion; molecular chaperone; stress-induced protein
 FEATURE
   1-33
   34-654
   654-646          $length 654; score 8; Da 1; Length 654;
                     $region aspargine-rich experimental; label NMT
 Query Match       0.7% ; score 8; Da 1; Length 654;
 Best Local Similarity 100.0%; Pred. No. 1,64e+01;
 Matches          8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db               595 LEBVARY 602
 OY              99 LEBVARY 106
 RESULT         11
 ENTRY          528103 $type fragment
 TITLE          Problematic DNA-directed DNA polymerase (EC 2.7.7.7) - g1111
 ORGANISM       Homo sapiens
 DATE           28-May-1993
 ACCESSIONS     GI:21-Feb-1997
 REFERENCES     821013
 AUTHORS        Robertson, M.H.; Royer, J.C.; Horzen, P.A.
 JOURNAL        Curr. Genet. [1991] 19:485-502
 REMARK        Homology between mitochondrial DNA of Agaricus bisporus and

```

[illegible]

```

#####
accession      528150
status         annotated
molcule-type  DNA
sequences      1..797 #label MOB
cross-references EMBL:J11313

GENETICS
genome         plasmid
keywords       DNA binding; allochondition; nucleotidytransferase
SMOVAR         length 797  checksum 1673

Query Match    0.74; Score 8; DB 2; Length 797;
Best Local Similarity 100.0%; Pred. No. 1,646-01;
Matches        0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 670 STARTSI 677
1.....1
Cy 569 STARTSI 976

RESULT 12
ENTRY
TITLE          histocytic polyphosphatase homology - human
AUTHOR        A. G. Olivieri, I. M. Olabe, I. J. Rallier, L. C.
REFERENCE      D.L. Lewis, R.A. Holmes, R.R. Neubauer, R.L.
               Nature (1992) 356:239-242
               The Lowe's oculocerebroretinal syndrome gene encodes a protein
               homologous to histocytic polyphosphatase 2 phosphatase
               #accession 629065
               #status preliminary
               #molcule-type mRNA
               #cross-references length 970  #molecule-weight 111772  #checksum 8909

SMOVAR         length 970  #molecule-weight 111772  #checksum 8909

Query Match    0.74; Score 8; DB 2; Length 970;
Best Local Similarity 100.0%; Pred. No. 1,646-01;
Matches        0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 332 LPPDVAR 33
1.....1
Cy 238 LPPDVAR 245

RESULT 13
ENTRY
SMOVAR         568440  #type complete
#####

```

[illegible]

RESULT 15
ENTRY J00182
TITLE hypothetical 3K protein (tm6-tpa15 intergenic region) -
ORGANISM Oryza sativa
DATE 31-Mar-1990
ACCESSION J00182
REFERENCE 100200
AUTHORS Shimada, H.; Wuttler, R.F.; Hixenbuttel, J.; Maeda, Y.; Hirai, A.; Sugita, M.
#sublation submitted to JFID, December 1989
#molecule-type DNA
#residues 1-39 #label SHI
#experimental_source cv. Nihonbare
KEYWORDS chloroplast
SUMMARY length 39 molecular-weight 4333 checksum 1192
Query Match 0.63; Score 7; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.32e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 5 SFLSBL 11
QY 366 SFLSBL 372

RESULT 16
ENTRY A49196
TITLE thyrotropin receptor - guinea pig (fragment)
ORGANISM guinea pig
DATE 19-Dec-1993
ACCESSION A49196
REFERENCE 100200
AUTHORS Roselli-Rehner, L.; Robbins, L.S.; Cone, R.D.
#journal Endocrinology (1992) 130:1851-1861
#title thyrotropin receptor cDNA encoding a nucleic acid 4x expressed
#cross-references MIM:9231878
#accession A49196
#molecule-type preliminary; not compared with conceptual translation
#residues 1-52 #label ROS
#experimental_source adipose tissue
#note sequence extracted from MIM backbone (MIM:89444)
#classification alpha-2-glycoprotein repeat homology

ACCESSIONS 14-Nov-1997
#molecule-type DNA
#residues 1-52 #label ROS
#experimental_source adipose tissue
#note sequence extracted from MIM backbone (MIM:89444)
#classification alpha-2-glycoprotein repeat homology

RESULT 17
ENTRY S01945
TITLE myosin alkali light chain (fragment)
ORGANISM mouse
DATE 03-Dec-1989
ACCESSION S01945
REFERENCE 100200
AUTHORS Cantor, R.; Sage, Y.; Litman, G.; Freeman, G.; Tung, J.S.; Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7360-7363
#journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7360-7363
#title myosin alkali light chain (fragment)
#cross-references MIM:860465
#molecule-type RNA
#residues 1-115 #label SHI
#experimental_source skeletal muscle
#note sequence extracted from MIM backbone (MIM:860465)
#classification myosin alkali light chain (fragment)
#accession S01945
#molecule-type RNA
#residues 1-115 #label SHI
#experimental_source skeletal muscle
#note sequence extracted from MIM backbone (MIM:860465)
#classification myosin alkali light chain (fragment)

SUMMARY length 52 checksum 4896
Query Match 0.63; Score 7; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 1.32e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 46 LKRYAN 52
QY 99 LKRYAN 105

RESULT 17
ENTRY S01945
TITLE myosin alkali light chain (fragment)
ORGANISM mouse
DATE 03-Dec-1989
ACCESSION S01945
REFERENCE 100200
AUTHORS Cantor, R.; Sage, Y.; Litman, G.; Freeman, G.; Tung, J.S.; Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7360-7363
#journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7360-7363
#title myosin alkali light chain (fragment)
#cross-references MIM:860465
#molecule-type RNA
#residues 1-115 #label SHI
#experimental_source skeletal muscle
#note sequence extracted from MIM backbone (MIM:860465)
#classification myosin alkali light chain (fragment)

ACCESSIONS 14-Nov-1997
#molecule-type DNA
#residues 1-52 #label ROS
#experimental_source adipose tissue
#note sequence extracted from MIM backbone (MIM:89444)
#classification alpha-2-glycoprotein repeat homology

RESULT 20
ENTRY A49196
TITLE thyrotropin receptor - guinea pig (fragment)
ORGANISM guinea pig
DATE 19-Dec-1993
ACCESSION A49196
REFERENCE 100200
AUTHORS Roselli-Rehner, L.; Robbins, L.S.; Cone, R.D.
#journal Endocrinology (1992) 130:1851-1861
#title thyrotropin receptor cDNA encoding a nucleic acid 4x expressed
#cross-references MIM:9231878
#accession A49196
#molecule-type preliminary; not compared with conceptual translation
#residues 1-52 #label ROS
#experimental_source adipose tissue
#note sequence extracted from MIM backbone (MIM:89444)
#classification alpha-2-glycoprotein repeat homology

DATE 25-Apr-1997 sequence_revision 25-Apr-1997 text_change
 ACCESSIONS 57066
 REFERENCE 57432
 AUTHORS Kaneo, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; Hirosewa, M.; Sugisue, M.; Murakami, Y.; Nakashita, K.; Nakashita, K.; Osumi, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Matsuda, A.; Tanaka, M.; Yasuda, M.; Tabata, S.
 JOURNAL DNA Res. (1996) 3:109-116
 TITLE The complete primary structure of the T-cell receptor gene from an allelically expressed cytotoxic human T-lymphocyte clone.
 cross-references EMBL:57066; FID:91001779; PDB:101111; PDB:91006607
 #molecule_type DNA
 #accession EMBL:57066; FID:91001779; PDB:101111; PDB:91006607
 #date 1996
 #organism Homo sapiens
 #library June 1996
 #length 138
 #molecular_weight 14443
 #checksum 2173

Query Match
 Best Local Similarity 100.0% Pred. No. 132+0-0
 Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 46 AVIEXOS 52
 611 AVIEXOS 817

RESULT 22
 TITLE PBM1A
 TYPE complete
 T-cell receptor alpha chain precursor V region (CTL-117) - human
 #formal_name Homo sapiens (common_name man)
 #date 1997
 #sequence_revision 17-Mar-1987
 #text_change
 #accessions A02016
 #reference A03750
 #authors Ledford, J.M.; Fraser, J.D.; Strominger, J.L.
 #journal J. Biol. Chem. (1986) 261:1171-1175
 #title The complete primary structure of the T-cell receptor gene from an allelically expressed cytotoxic human T-lymphocyte clone.
 #cross-references EMBL:86276770
 #accession A02016
 #molecule_type DNA
 #length 138
 #molecular_weight 14443
 #checksum 2173

Query Match
 Best Local Similarity 100.0% Pred. No. 132+0-0
 Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 58 LIPSLIC 64
 372 LIPSLIC 378

RESULT 24
 TITLE 564794
 TYPE complete
 hypothetical protein YL042c - yeast (Saccharomyces cerevisiae)
 #formal_name Saccharomyces cerevisiae
 #date 01-Aug-1995
 #sequence_revision 24-May-1996
 #text_change
 #accessions 564794
 #reference 564792
 #authors Wedler, H.; Wedler, E.; Scharfe, M.; Wandut, R.
 #journal J. Biol. Chem. (1995) 270:1151-1155
 #title The complete primary structure of the protein YL042c from the yeast Saccharomyces cerevisiae.
 #cross-references EMBL:573147; FID:91502316; PDB:245470; PDB:91502317; #accession A02016
 #molecule_type DNA
 #length 138
 #molecular_weight 14443
 #checksum 2173

GENETICS
 #experimental_source strain 5280C

SUMMARY
 #length 167
 #molecular_weight 19759
 #checksum 3500

Query Match
 Best Local Similarity 100.0% Pred. No. 132+0-0
 Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 36 DOLPHIN 42
 882 DOLPHIN 888

RESULT 25
 TITLE 504916
 TYPE fragment
 T-cell receptor delta chain precursor V-D-J region (clone RT08A) - human (fragments)
 #formal_name Homo sapiens (common_name man)
 #date 08-Sep-1997
 #sequence_revision 30-Sep-1997
 #text_change
 #accessions PLO078
 #reference PLO073
 #authors Takimaru, Y.; Reimann, J.; Michalopoulos, E.; Ciccone, E.; Mostoslavsky, L.; Maki, T.M.

COMMENT This sequence was derived from a human cytotoxic T-lymphocyte that is T37.74, T9.

GENETICS
 #gene CDB:TCRA
 #cross-references CDB:120404; OMIM:186880
 #position 1411.2-1411.2
 #description T-cell receptor alpha chain precursor V region (clone RT08A) - human (fragments)
 #date 07-Jul-1997
 #sequence_revision 29-Aug-1997
 #text_change
 #accessions J05476
 #reference J05476
 #authors Hishida, T.; Iwasaki, H.; Ishihara, K.; Shingawa, H.
 #journal J. Biol. Chem. (1996) 271:1151-1155
 #title The complete primary structure of the protein YL042c from the yeast Saccharomyces cerevisiae.
 #cross-references EMBL:573147; FID:91502316; PDB:245470; PDB:91502317; #accession A02016
 #molecule_type DNA
 #length 138
 #molecular_weight 14443
 #checksum 2173

Query Match
 Best Local Similarity 100.0% Pred. No. 132+0-0
 Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 4 LIGASVT 10
 697 LIGASVT 703

RESULT 23
 TITLE 510280
 TYPE complete
 hpcr protein Escherichia coli
 #formal_name Escherichia coli
 #date 30-Sep-1997
 #sequence_revision 30-Sep-1997
 #text_change
 #accessions S10279
 #reference S10279
 #authors Roper, D.L.; Fawcett, T.; Cooper, R.A.
 #journal Mol. Gen. Genet. (1997) 257:241-250
 #title The complete primary structure of the protein YL042c from the yeast Saccharomyces cerevisiae.
 #cross-references EMBL:573147; FID:91502316; PDB:245470; PDB:91502317; #accession A02016
 #molecule_type DNA
 #length 138
 #molecular_weight 14443
 #checksum 2173

GENETICS
 #experimental_source strain 5280C

SUMMARY
 #length 139
 #molecular_weight 15441
 #checksum 644

Query Match
 Best Local Similarity 100.0% Pred. No. 132+0-0
 Matches 7: Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 4 LIGASVT 10
 697 LIGASVT 703

RESULT 26
 TITLE 504916
 TYPE complete
 endocytosis-related gene - Pseudomonas aeruginosa
 #formal_name Pseudomonas aeruginosa
 #date 07-Jul-1997
 #sequence_revision 29-Aug-1997
 #text_change
 #accessions J05476
 #reference J05476
 #authors Hishida, T.; Iwasaki, H.; Ishihara, K.; Shingawa, H.
 #journal J. Biol. Chem. (1996) 271:1151-1155
 #title The complete primary structure of the protein YL042c from the yeast Saccharomyces cerevisiae.
 #cross-references EMBL:573147; FID:91502316; PDB:245470; PDB:91502317; #accession A02016
 #molecule_type DNA
 #length 138
 #molecular_weight 14443
 #checksum 2173

GENETICS
 #experimental_source strain 5280C

SUMMARY
 #length 139
 #molecular_weight 15441
 #checksum 644

Journal Nature (1997) 380:364-370
 Title The complete sequence of the hyperthermophilic sulfate-reducing archaeon *Archaeoglobus fulgidus*
 accession B69537
 status preliminary; nucleic acid sequence not shown;
 translation not shown
 #molecule_type DNA
 #residues 1-280 #label KLE
 #cross-reference GB:AB007821; FICR:AP2298
 #length 280 #molecular-weight 32856 #checksum 1366
 Query Match 0.64; Score 7; DB 2; Length 280;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 263 PFTLSS 268
 QY 365 PFTLSS 371
 RESULT 37
 ENTRY 521306
 TITLE Hypothetical protein - *Pseudomonas aeruginosa*
 ORIGIN 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
 DATE
 ACCESSIONS 521306
 #residues 221306
 #molecule_type DNA
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, September 1990
 #description DNA sequence analysis of the replication region of the *Pseudomonas aeruginosa* plasmid R91.5.
 #accession 521306 preliminary
 #molecule_type DNA
 #residues 1-796 #label DAV
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 290 #molecular-weight 32683 #checksum 5503
 Query Match 0.64; Score 7; DB 2; Length 296;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 253 STILDDQ 258
 QY 786 STILDDQ 792
 RESULT 38

REFERENCE 516681
 AUTHOR DeRubeis, V.; Kozlowski, J.H.
 JOURNAL Nucleic Acids Res. (1991) 19:1742
 TITLE Nucleotide sequence of a novel diverged human homeobox gene
 accession 516681
 status preliminary; nucleic acid sequence not shown;
 translation not shown
 #molecule_type mRNA
 #residues 1-316 #label DEC
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, January 1991
 #description DNA binding; homeobox; nucleus; transcription regulation
 #accession 516681 preliminary
 #molecule_type DNA
 #residues 1-316 #label ROX
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 316 #molecular-weight 34713 #checksum 1095
 Query Match 0.64; Score 7; DB 2; Length 316;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 267 QARPPH 273
 QY 205 QARPPH 211

RESULT 40
 ENTRY 519586
 TITLE Type complete
 ORIGIN 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
 DATE
 ACCESSIONS 519586
 #residues 519586
 #molecule_type DNA
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, January 1991
 #description DNA binding; homeobox; nucleus; transcription regulation
 #accession 519586 preliminary
 #molecule_type DNA
 #residues 1-316 #label ROX
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 316 #molecular-weight 34713 #checksum 1095
 Query Match 0.64; Score 7; DB 2; Length 316;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 267 QARPPH 273
 QY 205 QARPPH 211

ENTRY 515208
 TITLE Type complete
 ORIGIN 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
 DATE
 ACCESSIONS 515208
 #residues 515208
 #molecule_type DNA
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, January 1991
 #description DNA binding; homeobox; nucleus; transcription regulation
 #accession 515208 preliminary
 #molecule_type DNA
 #residues 1-316 #label ROX
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 316 #molecular-weight 34713 #checksum 1095
 Query Match 0.64; Score 7; DB 2; Length 316;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 267 QARPPH 273
 QY 205 QARPPH 211

REFERENCE 516681
 AUTHOR DeRubeis, V.; Kozlowski, J.H.
 JOURNAL Nucleic Acids Res. (1991) 19:1742
 TITLE Nucleotide sequence of a novel diverged human homeobox gene
 accession 516681
 status preliminary; nucleic acid sequence not shown;
 translation not shown
 #molecule_type mRNA
 #residues 1-316 #label DEC
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, January 1991
 #description DNA binding; homeobox; nucleus; transcription regulation
 #accession 516681 preliminary
 #molecule_type DNA
 #residues 1-316 #label ROX
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 316 #molecular-weight 34713 #checksum 1095
 Query Match 0.64; Score 7; DB 2; Length 316;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 267 QARPPH 273
 QY 205 QARPPH 211

RESULT 42
 ENTRY 569119
 TITLE Type complete
 ORIGIN 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 09-Sep-1997
 DATE
 ACCESSIONS 569119
 #residues 569119
 #molecule_type DNA
 #cross-reference EMBL:U00001; S.P.:Kishanlal, V.V.
 #status submitted to the EMBL Data Library, January 1991
 #description DNA binding; homeobox; nucleus; transcription regulation
 #accession 569119 preliminary
 #molecule_type DNA
 #residues 1-316 #label ROX
 #cross-reference EMBL:X54955; MID:g45410; PID:g45411
 #length 316 #molecular-weight 34713 #checksum 1095
 Query Match 0.64; Score 7; DB 2; Length 316;
 Best Local Similarity 100.04; Pred. No. 1.32e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 267 QARPPH 273
 QY 205 QARPPH 211

Accessions 556795
 REFERENCE 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Accessions 556795
 REFERENCE 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Accessions 556795
 REFERENCE 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Accessions 556795
 REFERENCE 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

Query Match 0.64; Score 7; DB 2; Length 347
 Best Local Similarity 100.0%; Pred. No. 1.32e+01
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 228 TALLMAN 234
 QY 1135 TALLMAN 1141

RESULT 47
 ENTRY 556795
 TITLE 556795
 ALTERNATE_NAMES 556795
 ORIGINISM 556795
 accession 556795
 #molecule_type DNA
 #cross-reference EMBL:449288; NID:g1006744; PID:g1006745; MIPS:Y10230c
 #note SCD:PET130
 #protein_location 19-201-1996
 #text_change

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

 Release 3.1a John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

DN DNA REPAIR: HYDROLASE, GLYCOSIDASE.
 AC ACT SITE 148 148 GENERAL BASE (BY SIMILARITY).
 DE SEQUENCE 305 AA: 34376 MW: 5050849 CMC32;
 Query Match 0.78; Score 8; DB 1; Length 305;
 Best Local Similarity 100.0%; Pred. No. 4,71e-02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 231 RYDORCE 238
 |||||||
 QY 105 RYDORCE 112

RESULT 4
 ID KMO5_MSV78 STANDARD: PRT: 342 AA.
 AC P10423;
 DE 01-OCT-1989 (REL. 10, CREATED)
 DE 01-OCT-1989 (REL. 11, LAST SEQUENCE UPDATE)
 DE 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE MOS SERINE/THROMBIN-PROTEIN KINASE TRANSFORMING PROTEIN (EC 2.7.1.-).
 CC V-MOS
 CC METEOROLIZATIVE SARCOMA VIRUS (CLOVE T8155)
 CC MOS SERINE/THROMBIN-PROTEIN KINASE TRANSFORMING PROTEIN (EC 2.7.1.-).
 CC OC OVERVIRAL
 CC [1]
 RN SEQUENCE FROM N.A.
 RA PRILEP, STOKING C., STACEY A., OSTERING W.;
 RA J. VIROL. 61:889-897(1987).
 CC -1- SIMILARITY WITH THE CONSERVED CATALYTIC DOMAIN OF SER/THR-
 CC KINASES
 CC PIR A16592; T8155; 11.
 CC PROTEIN KINASES
 CC DR PROSITE: P500107; PROTEIN KINASE, ATP, 1.
 CC DR PROSITE: P500108; PROTEIN KINASE, ATP, 1.
 CC DR TRANSFERASE: SERINE/THROMBIN-PROTEIN KINASE, ONCOGENE;
 CC ATP-BINDING;
 CC FT DOMAIN 63 338 PROTEIN KINASE
 CC FT NP_BIND 69 77 ATP (BY SIMILARITY)
 CC FT BINDING 135 256 CAMP
 CC FT ACT SITE 188 198 BY SIMILARITY
 CC SEQUENCE 342 AA: 37970 MW: 1662430 CMC32;
 Query Match 0.78; Score 8; DB 1; Length 342;
 Best Local Similarity 100.0%; Pred. No. 4,71e-02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 291 LAPSITCA 298
 |||||||
 QY 372 LAPSITCA 379

FT BINDING 334 334 CAMP.
 FT BINDING 343 343 CAMP.
 DE SEQUENCE 400 AA: 44962 MW: 17740099 CMC32;
 Query Match 0.78; Score 8; DB 1; Length 400;
 Best Local Similarity 100.0%; Pred. No. 4,71e-02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 106 REDDTR 113
 |||||||
 QY 461 REDDTR 468

RESULT 5
 ID KAP2_HMAN STANDARD: PRT: 403 AA.
 AC P13661; Q16823;
 DE 01-JAN-1990 (REL. 11, CREATED)
 DE 01-JAN-1990 (REL. 12, LAST SEQUENCE UPDATE)
 DE 01-SEP-1995 (REL. 31, LAST ANNOTATION UPDATE)
 DE CAMP-DEPENDENT PROTEIN KINASE TYPE II-ALPHA REGULATORY CHAIN.
 CC PRKAR2A OR PRK2 OR PRK2.
 CC HOMO SAPIENS (HUMAN).
 CC EMBL: X14588; C98448;
 CC EMBL: X14588; C98448;
 CC PIR: P05135; ZANK.
 CC SEQUENCE FROM N.A.
 RA TISSOT-DUPONT, H.;
 RA OTEN O.; WITTEBOST F.; SCOTT J.D.; HANSSON V.; JAHNSEN T.;
 RA PERS LETT. 246:57-64(1989).
 CC SEQUENCE OF 1-68 FROM N.A.
 DB 109 REDDTR 116
 |||||||
 QY 461 REDDTR 468

RESULT 5
 ID KAP2_BOVIN STANDARD: PRT: 400 AA.
 AC P05135;
 DE 21-SEP-1986 (REL. 01, CREATED)
 DE 01-SEP-1986 (REL. 12, LAST SEQUENCE UPDATE)
 DE 01-SEP-1986 (REL. 31, LAST ANNOTATION UPDATE)
 DE CAMP-DEPENDENT PROTEIN KINASE TYPE II-ALPHA REGULATORY CHAIN.
 CC PRKAR2A.
 CC BOS TAURUS (BOVINE).
 CC EMBL: X14588; C98448;
 CC EMBL: X14588; C98448;
 CC PIR: P05135; ZANK.
 CC SEQUENCE.
 RA TISSOT-DUPONT, H.;
 RA OTEN O.; WITTEBOST F.; SCOTT J.D.; HANSSON V.; JAHNSEN T.;
 RA PERS LETT. 246:57-64(1989).
 CC SEQUENCE OF 1-68 FROM N.A.
 DB 109 REDDTR 116
 |||||||
 QY 461 REDDTR 468

DB NEM: 1769310;
 DB PROSITE: P500888; CAMP_BINDING_1; 2.
 DB PROSITE: P500889; CAMP_BINDING_2; 2.
 DB CAMP-BINDING: PHOSPHORYLATION; DUBLIN; MULTIGENE FAMILY;
 CC ACETYLATION: 0
 CC FT INT_ACT 0
 CC FT DOMAIN 135 256 DIMERIZATION AND PHOSPHORYLATION.
 CC FT NP_BIND 260 403 CAMP.
 CC FT REPEAT 138 259
 CC FT REPEAT 260 403
 CC FT MOD_RES 91 98
 CC FT BINDING 207 207 CAMP.
 CC FT BINDING 216 216 CAMP.
 CC FT BINDING 317 317 CAMP.
 CC SEQUENCE 403 AA: 45387 MW: 96235283 CMC32;
 Query Match 0.78; Score 8; DB 1; Length 403;
 Best Local Similarity 100.0%; Pred. No. 4,71e-02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 109 REDDTR 116
 |||||||
 QY 461 REDDTR 468

[illegible]

07	3712	UNDELING	3718		
08	3713	UNDELING	3718		
09	3714	UNDELING	3718		
10	3715	UNDELING	3718		
11	3716	UNDELING	3718		
12	3717	UNDELING	3718		
13	3718	UNDELING	3718		
14	3719	UNDELING	3718		
15	3720	UNDELING	3718		
16	3721	UNDELING	3718		
17	3722	UNDELING	3718		
18	3723	UNDELING	3718		
19	3724	UNDELING	3718		
20	3725	UNDELING	3718		
21	3726	UNDELING	3718		
22	3727	UNDELING	3718		
23	3728	UNDELING	3718		
24	3729	UNDELING	3718		
25	3730	UNDELING	3718		
26	3731	UNDELING	3718		
27	3732	UNDELING	3718		
28	3733	UNDELING	3718		
29	3734	UNDELING	3718		
30	3735	UNDELING	3718		
31	3736	UNDELING	3718		
32	3737	UNDELING	3718		
33	3738	UNDELING	3718		
34	3739	UNDELING	3718		
35	3740	UNDELING	3718		
36	3741	UNDELING	3718		
37	3742	UNDELING	3718		
38	3743	UNDELING	3718		
39	3744	UNDELING	3718		
40	3745	UNDELING	3718		
41	3746	UNDELING	3718		
42	3747	UNDELING	3718		
43	3748	UNDELING	3718		
44	3749	UNDELING	3718		
45	3750	UNDELING	3718		
46	3751	UNDELING	3718		
47	3752	UNDELING	3718		
48	3753	UNDELING	3718		
49	3754	UNDELING	3718		
50	3755	UNDELING	3718		
51	3756	UNDELING	3718		
52	3757	UNDELING	3718		
53	3758	UNDELING	3718		
54	3759	UNDELING	3718		
55	3760	UNDELING	3718		
56	3761	UNDELING	3718		
57	3762	UNDELING	3718		
58	3763	UNDELING	3718		
59	3764	UNDELING	3718		
60	3765	UNDELING	3718		
61	3766	UNDELING	3718		
62	3767	UNDELING	3718		
63	3768	UNDELING	3718		
64	3769	UNDELING	3718		
65	3770	UNDELING	3718		
66	3771	UNDELING	3718		
67	3772	UNDELING	3718		
68	3773	UNDELING	3718		
69	3774	UNDELING	3718		
70	3775	UNDELING	3718		
71	3776	UNDELING	3718		
72	3777	UNDELING	3718		
73	3778	UNDELING	3718		
74	3779	UNDELING	3718		
75	3780	UNDELING	3718		
76	3781	UNDELING	3718		
77	3782	UNDELING	3718		
78	3783	UNDELING	3718		
79	3784	UNDELING	3718		
80	3785	UNDELING	3718		
81	3786	UNDELING	3718		
82	3787	UNDELING	3718		
83	3788	UNDELING	3718		
84	3789	UNDELING	3718		
85	3790	UNDELING	3718		
86	3791	UNDELING	3718		
87	3792	UNDELING	3718		
88	3793	UNDELING	3718		

[illegible]

Dh	77 ARGENS 03
Oy	243 ARGENS 249
RESULT 19	
ID	RVCV_PSEAE STANDARD: PRT; 174 NA.
G1	01-NOV-1987 (REL. 35, CREATED)
D7	01-NOV-1987 (REL. 35, LAST SEQUENCE UPDATE)
D7	01-NOV-1987 (REL. 35, LAST SEQUENCE UPDATE)
D7	01-NOV-1987 (REL. 35, LAST SEQUENCE UPDATE)
D7	CROSSOVER JUNCTION PHOSPHOTRANSFERASE GENE (EC 3.1.22.4) (HOLLAND JUNCTION RESOLUASE GENE).
G3	RVCV
OC	PEDOMONAS AEROGIENSA.
OC	PERIDIUMMAGNIFICA.
OC	PERIDIUMMAGNIFICA: SCOTOBACTERIA, AEROBIC RODS AND COCCI;
RP	SEQUENCE FROM N.A.
RP	SEQUENCE FROM N.A.
PC	STAIN=PAOI.
PC	MUTATION: 9713681
RL	GENE 182,657-9139961.
CC	-1- FUNCTION: NUCLEASE THAT RESOLVES HOLLAND JUNCTION INTERMEDIATES IN GENETIC RECOMBINATION. CLEAVES THE CRUCIFORM STRUCTURE IN AT SITES SYMMETRICALLY OPPOSED AT THE JUNCTION IN THE HOMOLOGOUS ARMS AND LEAVES A 5'TERMINAL PHOSPHATE AND A 3'TERMINAL HYDROXYL GROUP (BY SHIMAMURA).
CC	-1- CATALYTIC ACTIVITY: PHOSPHOTRANSFERASE N.A. JUNCTION SUCH AS CROSSOVER JUNCTIONS, PHOSPHOTRANSFERASE N.A. JUNCTIONS BETWEEN NONHOMOLOGOUS DNA
CC	DUPLEXES (HOLLAND JUNCTION), CROSSOVER BETWEEN NONHOMOLOGOUS DNA
CC	REPAIR: DNA REPAIR: DNA RECOMBINATION.
CC	EMBL: D81188, G131840.
CC	SEQUENCE 174 NA: 16556 NM: 16556-635 CMC12.
Qy	Query Match 0.68; Score 7; db 1; Length 174; Best Local Similarity 100.0%; Pred. No. 5, 9e+00; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Dh	164 GARRGG 170
Oy	242 GARRGG 248
RESULT 20	
ID	TOPY_YERNH STANDARD: PRT; 182 NA.
G1	07-APR-1989 (REL. 33, CREATED)
D7	01-AUG-1993 (REL. 33, LAST SEQUENCE UPDATE)
D7	01-JUL-1993 (REL. 26, LAST ANNOTATION UPDATE)

[illegible]

RN	(1)
RP	SEQUENCE FROM N.A.
RA	MEHLER, J., KATZ, D., AND R. H. MULLIS, 1987, PNAS 84:1318-1321.
RX	MEDLINE 95194318.
RA	ARITA M., SAND T.E., MIYATA A., TANABE T., TAKAHASHI E.,
RA	BARDEN H.J., ABBI E., INOUE K.,
RL	J. AM. CHEM. SOC. 113:406-437(1991).
RP	SEQUENCE FROM N.A.
RA	DENG H.X., HERBASTI N., SIDDIQIOU T.:
RL	SUBMITTED (FEB-1995) TO EMBL/GENBANK/DDDB DATA BANKS.
RP	VALIANT AYED GJM-101.
RA	MEHLER, 96036648.
RA	GORDON T., ARITA M., ABBI E., INOUE K., YOKOTA T., FUKUI Y.,
RA	OGURA M., MED. SCI. 313:1111-1118(1995).
CC	-1- FUNCTION: RIBOS ALPHA-TGPOPHROL AND ENHANCES ITS TRANSFER BETWEEN
CC	SEPARATE MEMBRANES.
CC	-1- DISCELLULAR LOCATION: CYTOSOLIC.
CC	-1- DISEASE: SPINocerebellar ATROPHY. CLINICAL COURSE OF ALXIA WITH ISOLATED
CC	VITAMIN E DEFICIENCY (AVED). AN AUTOSOMAL RECESSIVE DISEASE
CC	CHARACTERIZED BY SPINocerebellar DEGENERATION.
CC	-1- SIMILARITY: TO CELLULAR RETINALDEHYDE-BINDING PROTEIN.
DR	EMBL, U21938. C126182; -
DR	KIM, 600415; -
DR	MIN, 277460; -
DM	TRANSFERRIN: DISEASE MUTATION. H -> Q (IN AVED).
PT	AVED, 501
FT	CONFLICT 271
FT	CONFLICT 271
SO	SEQUENCE 278 AA; 31749 MW; B3073797 CRC12;
Query March 0 6%	
Best local similarity 100.0%; Pred. No. 5 gsta=00;	
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Oy	230 RENOVTL 236
RESULT: 37	
ID	PARK STROC STANDARD: PRT: 278 AA.
AC	P43351.
DT	01-NOV-1995 (REL. 32, CREATED)
DT	01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT	01-NOV-1995 (REL. 32, FIRST SEQUENCE UPDATE)
DE	23 TO 26 MEMBRANE PROTEIN IN PSA 5 REGION (ORF1).
DE	23 TO 26 MEMBRANE PROTEIN IN PSA 5 REGION (ORF1).
CC	STRUTPOCOCUS GORDONI CHALLIS.
CC	PROBABLYA. FRINCIGUES; CCCT1; STRUTPOCOCALAE.

[illegible][illegible]


```

Db          QY      97 GSTRND 103
                |||||
QY      1116 GSTRND 1122

RESULT      34
ID     PT30_YEAST    STANDARD:      PRT:   347 AA.
AC     R47085-1986 (REL. 33, CREATD)
DT     01-FEB-1996 (REL. 33), LAST SEQUENCE UPDATE)
DE     PT10-1996 (REL. 34, LAST ANNOTATION UPDATE)
PE     PT110 PROTEIN.
OR     PERCUTANEOUS GLYCEROL KINASE OR A1283
OS     NEUROSPORA CRASSISPORA BAKER'S YEAST)
OC     EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOMETECES.
LN     [1]
RF     SEQUENCE FROM N.A.
RP     SUBMITTED (OCT-1995) TO EMBL/GENBANK/DDJJ DATA BANKS.
RL     [2]
RR     [2]
RS     SEQUENCE FROM N.A.
RP     SUBMITTED (SEP-1995) TO EMBL/GENBANK/DDJJ DATA BANKS.
RL     [2]
RR     [2]
RC     FUNCTION: NOT KNOWN. PT110 MUTANTS ARE DEFICIENT IN MITOCHONDRIAL
CC     - PROTEIN SYNTHESIS.
CD     - POTENTIAL LOCATION: MITOCHONDRIAL MATRIX (POTENTIAL).
CE     DBL: J11717.G.DJ006745; -.
DR     EMBL: J11717.G.DJ006745; -.
DS     SCD: L0003025; PT110.
FM     MITOCHONDRION.
SQ     SEQUENCE 347 AA; 38612 MW; 65A9097F CRC32;.

Query Match      0.6% Score 7; Db 1; Length 347;
Best Local Similarity 100.0%; Pred. No 5,98e+00;
Matches       7; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

Db          QY      67 STDND 73
                |||||
QY      786 STDND 792

RESULT      35
ID     KROGALICOR    STANDARD:      PRT:   348 AA.
AC     D03083-1997 (REL. 35, CREATD)
DT     01-NOV-1997 (REL. 35), LAST SEQUENCE UPDATE)
DE     01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DS     UDP-GALACTOSE 4-EPIMERASE (EC 5.1.3.2) (UDP-GALACTOSE 4-EPIMERASE2)

```

[illegible][illegible][illegible]

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

764	5	0	5	730	5	0	5	018520	REVERSE TRANSCRIPTASE
765	5	0	5	731	4	0	1	014380	NOTCH SIGNALING FACTOR RE
766	5	0	5	733	5	0	0	024250	ADENOSINE DEAMINASE
767	5	0	5	735	5	0	0	024250	MATERN PROTEIN PRECURS
768	5	0	5	737	14	0	1	018881	CASIN PROTEIN VPL VPL
769	5	0	5	739	5	0	0	018820	COENED FOR BY C. ERYTH
770	5	0	5	742	11	0	0	008658	NUCLEOSIDE-NUCLEIC
771	5	0	5	749	4	0	1	013554	NUCLEOSIDE-NUCLEIC
772	5	0	5	750	4	0	0	013118	SEAPROHIN V.
773	5	0	5	752	14	0	0	003128	COAT PROTEIN VPL (CONT
774	5	0	5	751	14	0	0	003128	COAT PROTEIN VPL (CONT
775	5	0	5	753	5	0	0	003845	P-TANSPORIN (CLONED JB
776	5	0	5	757	2	0	0	005952	D-(-)-3-HYDROXYBUTYL
777	5	0	5	759	14	0	0	007897	POLYMERASE 2
778	5	0	5	759	14	0	0	007897	POLYMERASE 2
779	5	0	5	760	14	0	0	011108	CASIN F66111 PROTEIN BAS
780	5	0	5	764	3	0	1	042376	CHEMOTHERICAL 85.4 ND P
781	5	0	5	765	3	0	1	042376	CHEMOTHERICAL 85.4 ND P
782	5	0	5	767	3	0	1	041800	HYDROLYTICAL 85.1 ND P
783	5	0	5	771	1	0	0	014100	HYDROLYTICAL 85.1 ND P
784	5	0	5	776	14	0	0	008191	VPA PROTEIN
785	5	0	5	776	14	0	0	008191	VPA PROTEIN
786	5	0	5	776	14	0	0	008191	VPA PROTEIN
787	5	0	5	776	14	0	0	008191	VPA PROTEIN
788	5	0	5	776	14	0	0	008191	VPA PROTEIN
789	5	0	5	777	14	0	0	012193	OTHER CASIN PROTEIN V
790	5	0	5	780	13	0	0	0190424	OTHER CASIN PROTEIN V
791	5	0	5	780	11	0	0	008178	CD44 PROTEIN
792	5	0	5	780	11	0	0	008178	CD44 PROTEIN
793	5	0	5	781	13	0	0	006186	PLATELET-RECEPTOR
794	5	0	5	785	6	0	5	018643	INTRON OLF.
795	5	0	5	786	6	0	5	018643	INTRON OLF.
796	5	0	5	789	14	0	0	006547	RNA HELICASE-PRIMASE C
797	5	0	5	786	11	0	0	0141390	PHOSPHATASE GROWTH PACT
798	5	0	5	800	4	0	0	0090505	PHOSPHATASE GROWTH PACT
799	5	0	5	800	4	0	0	0090505	PHOSPHATASE GROWTH PACT
800	5	0	5	800	10	0	0	0493128	SIMILAR TO DISSE RES
801	5	0	5	800	10	0	0	0493128	SIMILAR TO DISSE RES
802	5	0	5	802	10	0	0	013045	ROOT HAIR GROWTH FACT
803	5	0	5	802	10	0	0	013045	ROOT HAIR GROWTH FACT
804	5	0	5	802	10	0	0	013045	ROOT HAIR GROWTH FACT
805	5	0	5	803	4	0	1	013112	TO DENDRODENDR/ACET
806	5	0	5	806	1	0	0	017218	TO DENDRODENDR/ACET
807	5	0	5	807	3	0	0	006479	CHROMOSOME XII COND
808	5	0	5	807	3	0	0	006479	CHROMOSOME XII COND
809	5	0	5	808	3	0	0	017783	FROM BASHES 105076 TO
810	5	0	5	810	10	0	0	0123446	SALT INDUCIBLE PROTEIN
811	5	0	5	810	10	0	0	0123446	SALT INDUCIBLE PROTEIN
812	5	0	5	816	11	0	0	019474	NEUROCAN (FRAGMENT).
813	5	0	5	816	11	0	0	019474	NEUROCAN (FRAGMENT).
814	5	0	5	818	14	0	1	099311	92K PROTEIN.

815	6	0	5	818	13	0.17152	FIBROBLAST GROWTH FACT
816	6	0	5	819	4	0.11705	MENA (ELN/OSF1) FOR OR
817	6	0	5	820	4	0.41188	5% SHMP 100 KD PROTEIN
818	6	0	5	821	5	0.61229	BETA-CATENIN
819	6	0	5	822	1	0.27154	CONSERVED PROTEIN
820	6	0	5	823	1	0.27154	CONSERVED PROTEIN
821	6	0	5	824	1	0.19173	FIBROBLAST GROWTH FACT
822	6	0	5	825	1	0.12028	FIBROBLAST GROWTH FACT
823	6	0	5	826	1	0.27154	FIBROBLAST GROWTH FACT
824	6	0	5	827	13	0.65480	CASPD PROTEIN (FRAHE
825	6	0	5	828	14	0.50131	100K PROTEIN
826	6	0	5	829	2	0.55035	CLAC
827	6	0	5	830	2	0.41023	CLAC
828	6	0	5	831	2	0.41023	CLAC
829	6	0	5	832	1	0.51537	CYTIC INDUCTORIDE-GATZ
830	6	0	5	833	5	0.51537	CYTIC INDUCTORIDE-GATZ
831	6	0	5	834	5	0.860	HYPOPHYSICAL 99.0 KD P
832	6	0	5	835	5	0.860	HYPOPHYSICAL 99.0 KD P
833	6	0	5	836	5	0.860	HYPOPHYSICAL 99.0 KD P
834	6	0	5	837	5	0.860	HYPOPHYSICAL 99.0 KD P
835	6	0	5	838	5	0.860	HYPOPHYSICAL 99.0 KD P
836	6	0	5	839	5	0.860	HYPOPHYSICAL 99.0 KD P
837	6	0	5	840	5	0.860	HYPOPHYSICAL 99.0 KD P
838	6	0	5	841	5	0.860	HYPOPHYSICAL 99.0 KD P
839	6	0	5	842	5	0.860	HYPOPHYSICAL 99.0 KD P
840	6	0	5	843	5	0.860	HYPOPHYSICAL 99.0 KD P
841	6	0	5	844	5	0.860	HYPOPHYSICAL 99.0 KD P
842	6	0	5	845	5	0.860	HYPOPHYSICAL 99.0 KD P
843	6	0	5	846	5	0.860	HYPOPHYSICAL 99.0 KD P
844	6	0	5	847	5	0.860	HYPOPHYSICAL 99.0 KD P
845	6	0	5	848	5	0.860	HYPOPHYSICAL 99.0 KD P
846	6	0	5	849	5	0.860	HYPOPHYSICAL 99.0 KD P
847	6	0	5	850	5	0.860	HYPOPHYSICAL 99.0 KD P
848	6	0	5	851	5	0.860	HYPOPHYSICAL 99.0 KD P
849	6	0	5	852	5	0.860	HYPOPHYSICAL 99.0 KD P
850	6	0	5	853	5	0.860	HYPOPHYSICAL 99.0 KD P
851	6	0	5	854	5	0.860	HYPOPHYSICAL 99.0 KD P
852	6	0	5	855	5	0.860	HYPOPHYSICAL 99.0 KD P
853	6	0	5	856	5	0.860	HYPOPHYSICAL 99.0 KD P
854	6	0	5	857	5	0.860	HYPOPHYSICAL 99.0 KD P
855	6	0	5	858	5	0.860	HYPOPHYSICAL 99.0 KD P
856	6	0	5	859	5	0.860	HYPOPHYSICAL 99.0 KD P
857	6	0	5	860	5	0.860	HYPOPHYSICAL 99.0 KD P
858	6	0	5	861	5	0.860	HYPOPHYSICAL 99.0 KD P
859	6	0	5	862	5	0.860	HYPOPHYSICAL 99.0 KD P
860	6	0	5	863	5	0.860	HYPOPHYSICAL 99.0 KD P
861	6	0	5	864	5	0.860	HYPOPHYSICAL 99.0 KD P
862	6	0	5	865	5	0.860	HYPOPHYSICAL 99.0 KD P
863	6	0	5	866	5	0.860	HYPOPHYSICAL 99.0 KD P
864	6	0	5	867	5	0.860	HYPOPHYSICAL 99.0 KD P
865	6	0	5	868	5	0.860	HYPOPHYSICAL 99.0 KD P
866	6	0	5	869	5	0.860	HYPOPHYSICAL 99.0 KD P
867	6	0	5	870	5	0.860	HYPOPHYSICAL 99.0 KD P
868	6	0	5	871	5	0.860	HYPOPHYSICAL 99.0 KD P
869	6	0	5	872	5	0.860	HYPOPHYSICAL 99.0 KD P
870	6	0	5	873	5	0.860	HYPOPHYSICAL 99.0 KD P
871	6	0	5	874	5	0.860	HYPOPHYSICAL 99.0 KD P
872	6	0	5				

866	5	0	5	1051	5	01801	MD1 TRANSFERROGENS	6.91e+00
867	5	0	5	1051	5	011024	CASID V2	6.91e+00
868	5	0	5	1053	1	065593	POLY-POLYMERIN	6.91e+00
869	5	0	5	1061	10	048487	BR-LIPOXGENASE	6.91e+00
870	5	0	5	1066	5	016025	BR-LIPOXGENASE-ALLELE	6.91e+00
871	5	0	5	1072	5	045532	COL-1 S1/BLKX YINOS	6.91e+00
872	5	0	5	1072	5	045532	COL-1 S1/BLKX YINOS	6.91e+00
873	5	0	5	1072	11	035482	HIGH MOLECULAR-WEIGHT	6.91e+00
874	5	0	5	1080	2	048378	LACTA PROTEIN	6.91e+00
875	5	0	5	1089	4	035252	PROTEIN BINDING PROTEIN	6.91e+00
876	5	0	5	1093	14	039868	HEV-6	6.91e+00
877	5	0	5	1131	14	039868	HEV-6	6.91e+00
878	5	0	5	1136	5	019371	PLIF-6	6.91e+00
879	5	0	5	1136	5	048582	ALINOT-2R	6.91e+00
880	5	0	5	1136	5	048582	ALINOT-2R	6.91e+00
881	5	0	5	1133	5	045632	HN907.2R	6.91e+00
882	5	0	5	1134	10	049433	HYPOCRETIN-137.3 KD	6.91e+00
883	5	0	5	1134	10	055318	POTATIVE RECEPTOR PROT	6.91e+00
884	5	0	5	1136	5	023465	SMITHSONIAN NUCLEOPROIN	6.91e+00
885	5	0	5	1166	5	023465	SMITHSONIAN NUCLEOPROIN	6.91e+00
886	5	0	5	1168	2	P74359	HYPOCRETIN-137.3 KD	6.91e+00
887	5	0	5	1168	2	030631	HYPOCRETIN-137.3 KD	6.91e+00
888	5	0	5	1175	11	031526	PROTEIN-135	6.91e+00
889	5	0	5	1182	4	039495	DEPLA	6.91e+00
890	5	0	5	1182	4	039495	DEPLA	6.91e+00
891	5	0	5	1182	4	039495	DEPLA	6.91e+00
892	5	0	5	1186	14	013101	ALPHA-1-VOLING-DEPEND	6.91e+00
893	5	0	5	1186	14	013101	ALPHA-1-VOLING-DEPEND	6.91e+00
894	5	0	5	1200	2	P73340	GLONERIN EPITHELIAL	6.91e+00
895	5	0	5	1212	4	060105	CHROMOSOME SEGREGATION	6.91e+00
896	5	0	5	1216	4	016182	KIA00559 PROTEIN (PRAC	6.91e+00
897	5	0	5	1227	5	P06091	PROTEIN T135	6.91e+00
898	5	0	5	1232	5	P06091	PROTEIN T135	6.91e+00
899	5	0	5	1230	11	P97336	MELICANSE	6.91e+00
900	5	0	5	1235	5	024135	T12120	6.91e+00
901	5	0	5	1235	5	024135	T12120	6.91e+00
902	5	0	5	1245	2	045512	TIROSINE KINASE (PRAC	6.91e+00
903	5	0	5	1245	2	045512	DELTA ENOXONIN (PRAC	6.91e+00
904	5	0	5	1245	2	045512	DELTA ENOXONIN (PRAC	6.91e+00
905	5	0	5	1262	5	024112	P-185 PROTO-ONCOGENE P	6.91e+00
906	5	0	5	1262	5	024112	STW-B	6.91e+00
907	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
908	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
909	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
910	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
911	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
912	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
913	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
914	5	0	5	1265	4	092833	DEBROGAL-PROTEIN	6.91e+00
915	5	0	5	1374	14	P99442	MAJOR CASID PROTEIN	6.91e+00
916	5	0	5	1375	14	039383	COMTEMPART OF HSV-1 G	6.91e+00

317	6	0	5	1419	11	063113	KLAF58B ALPHA 1 TYPE
318	6	0	5	1466	44	063014	KLAF58B ALPHA 2 TYPE
319	6	0	5	1441	41	051074	KLAF58B ALPHA 3 TYPE
320	6	0	5	1441	5	028453	KLAF58B ALPHA 4 TYPE
321	6	0	5	1442	11	063013	KLAF58B ALPHA 5 TYPE
322	6	0	5	1442	11	063013	KLAF58B ALPHA 6 TYPE
323	6	0	5	1401	11	031501	KLAF58B ALPHA 7 TYPE
324	6	0	5	1402	11	031501	KLAF58B ALPHA 8 TYPE
325	6	0	5	1402	11	031501	KLAF58B ALPHA 9 TYPE
326	6	0	5	1402	11	031501	KLAF58B ALPHA 10 TYPE
327	6	0	5	1402	11	031501	KLAF58B ALPHA 11 TYPE
328	6	0	5	1402	11	031501	KLAF58B ALPHA 12 TYPE
329	6	0	5	1402	11	031501	KLAF58B ALPHA 13 TYPE
330	6	0	5	1402	11	031501	KLAF58B ALPHA 14 TYPE
331	6	0	5	1402	11	031501	KLAF58B ALPHA 15 TYPE
332	6	0	5	1402	11	031501	KLAF58B ALPHA 16 TYPE
333	6	0	5	1402	11	031501	KLAF58B ALPHA 17 TYPE
334	6	0	5	1402	11	031501	KLAF58B ALPHA 18 TYPE
335	6	0	5	1402	11	031501	KLAF58B ALPHA 19 TYPE
336	6	0	5	1402	11	031501	KLAF58B ALPHA 20 TYPE
337	6	0	5	1402	11	031501	KLAF58B ALPHA 21 TYPE
338	6	0	5	1402	11	031501	KLAF58B ALPHA 22 TYPE
339	6	0	5	1402	11	031501	KLAF58B ALPHA 23 TYPE
340	6	0	5	1402	11	031501	KLAF58B ALPHA 24 TYPE
341	6	0	5	1402	11	031501	KLAF58B ALPHA 25 TYPE
342	6	0	5	1402	11	031501	KLAF58B ALPHA 26 TYPE
343	6	0	5	1402	11	031501	KLAF58B ALPHA 27 TYPE
344	6	0	5	1402	11	031501	KLAF58B ALPHA 28 TYPE
345	6	0	5	1402	11	031501	KLAF58B ALPHA 29 TYPE
346	6	0	5	1402	11	031501	KLAF58B ALPHA 30 TYPE
347	6	0	5	1402	11	031501	KLAF58B ALPHA 31 TYPE
348	6	0	5	1402	11	031501	KLAF58B ALPHA 32 TYPE
349	6	0	5	1402	11	031501	KLAF58B ALPHA 33 TYPE
350	6	0	5	1402	11	031501	KLAF58B ALPHA 34 TYPE
351	6	0	5	1402	11	031501	KLAF58B ALPHA 35 TYPE
352	6	0	5	1402	11	031501	KLAF58B ALPHA 36 TYPE
353	6	0	5	1402	11	031501	KLAF58B ALPHA 37 TYPE
354	6	0	5	1402	11	031501	KLAF58B ALPHA 38 TYPE
355	6	0	5	1402	11	031501	KLAF58B ALPHA 39 TYPE
356	6	0	5	1402	11	031501	KLAF58B ALPHA 40 TYPE
357	6	0	5	1402	11	031501	KLAF58B ALPHA 41 TYPE
358	6	0	5	1402	11	031501	KLAF58B ALPHA 42 TYPE
359	6	0	5	1402	11	031501	KLAF58B ALPHA 43 TYPE
360	6	0	5	1402	11	031501	KLAF58B ALPHA 44 TYPE
361	6	0	5	1402	11	031501	KLAF58B ALPHA 45 TYPE
362	6	0	5	1402	11	031501	KLAF58B ALPHA 46 TYPE
363	6	0	5	1402	11	031501	KLAF58B ALPHA 47 TYPE
364	6	0	5	1402	11	031501	KLAF58B ALPHA 48 TYPE
365	6	0	5	1402	11	031501	KLAF58B ALPHA 49 TYPE
366	6	0	5	1402	11	031501	KLAF58B ALPHA 50 TYPE
367	6	0	5	1402	11	031501	KLAF58B ALPHA 51 TYPE

[illegible]

```

RESULT 17          PRELIMINARY;      PRT;    869 MA.
ID   Q34144
SV   1
DT   01-NOV-1996 (TRENDRELE. 01, CREATED)
DT   01-NOV-1996 (TRENDRELE. 01, LAST SEQUENCE UPDATE)
DT   01-JAN-1998 (TRENDRELE. 05, LAST ANNOTATION UPDATE)
ID   UNKNNED PROTEIN
CM   1#
OS   ENDOPLASMIC RETICULUM; ARTHROPODA; INSECTA; DIPTERA.
OC   EMBRYONAL DEVELOPMENT; FERTILIZATION; Oogenesis.
ID   PHOSPHATASE
SV   1
RA   PHOSPHATASE FROM H.A.
RP   SEQUENCE FROM H.A.
PR   PARK W.J., LIU J., SHARP R.J., ADLER P.N.;
     NBER, 1994; 1255002(11965).
DN   DBEML:0673281; GI:1255002(11965).
DR   PHYBASE: Phy00001259; I.G.
SO   SEQUENCE    869 AA;  9785 MW;  E35B8B88 CRC32;
Query Match
Best Local Similarity 100.0% ; Score 8 ; DB 5; Length 869;
Matches      8; Complementary      0; Mismatches      0; Indels      0; Gaps
DB       410 SEQSTAT; 437
ID       251 SENSEMAP; 258

```

Mon Dec 21 09:28:30 1998

US-08-951-733-20.rsp

Page 35

[illegible]

```

RC STRAIN-D-2:
RC MEDLINE: 9519113.
RA HADDAD, L., GANTZER M. C., LE CAER J. P., DE LODRESSE N., SPERLING L.;
RA GENBANK: 663639.1995.
RA (2)
RA SEQUENCE FROM N. A.
RC STRAIN-D-2:
RA HADDAD, L., GANTZER M. C., LE CAER J. P., DE LODRESSE N., SPERLING L.;
RA MOL. BIOL. CELL. 6:649-659.1995.
RA EMBL: 027510. 0881410: -.
RA MATRIX PROTEIN.
RA NON-TER 1
DB SEQUENCE 23 AA; 2706 MW; A87C0F4 CRC32;
Query Match 0.64; Score 7; DB 10; Length 37;
DB Local Similarity 100.04; Pred. No 1.47e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps
DB 7 VSTLND 13
CY 785 VSTLND 791
RESOUZ 23
AC 065832:
AC 065833:
AC 065834:
AC 065835:
AC 065836:
AC 065837:
AC 065838:
AC 065839:
AC 065840:
AC 065841:
AC 065842:
AC 065843:
AC 065844:
AC 065845:
AC 065846:
AC 065847:
AC 065848:
AC 065849:
AC 065850:
AC 065851:
AC 065852:
AC 065853:
AC 065854:
AC 065855:
AC 065856:
AC 065857:
AC 065858:
AC 065859:
AC 065860:
AC 065861:
AC 065862:
AC 065863:
AC 065864:
AC 065865:
AC 065866:
AC 065867:
AC 065868:
AC 065869:
AC 065870:
AC 065871:
AC 065872:
AC 065873:
AC 065874:
AC 065875:
AC 065876:
AC 065877:
AC 065878:
AC 065879:
AC 065880:
AC 065881:
AC 065882:
AC 065883:
AC 065884:
AC 065885:
AC 065886:
AC 065887:
AC 065888:
AC 065889:
AC 065890:
AC 065891:
AC 065892:
AC 065893:
AC 065894:
AC 065895:
AC 065896:
AC 065897:
AC 065898:
AC 065899:
AC 065900:
AC 065901:
AC 065902:
AC 065903:
AC 065904:
AC 065905:
AC 065906:
AC 065907:
AC 065908:
AC 065909:
AC 065910:
AC 065911:
AC 065912:
AC 065913:
AC 065914:
AC 065915:
AC 065916:
AC 065917:
AC 065918:
AC 065919:
AC 065920:
AC 065921:
AC 065922:
AC 065923:
AC 065924:
AC 065925:
AC 065926:
AC 065927:
AC 065928:
AC 065929:
AC 065930:
AC 065931:
AC 065932:
AC 065933:
AC 065934:
AC 065935:
AC 065936:
AC 065937:
AC 065938:
AC 065939:
AC 065940:
AC 065941:
AC 065942:
AC 065943:
AC 065944:
AC 065945:
AC 065946:
AC 065947:
AC 065948:
AC 065949:
AC 065950:
AC 065951:
AC 065952:
AC 065953:
AC 065954:
AC 065955:
AC 065956:
AC 065957:
AC 065958:
AC 065959:
AC 065960:
AC 065961:
AC 065962:
AC 065963:
AC 065964:
AC 065965:
AC 065966:
AC 065967:
AC 065968:
AC 065969:
AC 065970:
AC 065971:
AC 065972:
AC 065973:
AC 065974:
AC 065975:
AC 065976:
AC 065977:
AC 065978:
AC 065979:
AC 065980:
AC 065981:
AC 065982:
AC 065983:
AC 065984:
AC 065985:
AC 065986:
AC 065987:
AC 065988:
AC 065989:
AC 065990:
AC 065991:
AC 065992:
AC 065993:
AC 065994:
AC 065995:
AC 065996:
AC 065997:
AC 065998:
AC 065999:
AC 066000:
AC 066001:
AC 066002:
AC 066003:
AC 066004:
AC 066005:
AC 066006:
AC 066007:
AC 066008:
AC 066009:
AC 066010:
AC 066011:
AC 066012:
AC 066013:
AC 066014:
AC 066015:
AC 066016:
AC 066017:
AC 066018:
AC 066019:
AC 066020:
AC 066021:
AC 066022:
AC 066023:
AC 066024:
AC 066025:
AC 066026:
AC 066027:
AC 066028:
AC 066029:
AC 066030:
AC 066031:
AC 066032:
AC 066033:
AC 066034:
AC 066035:
AC 066036:
AC 066037:
AC 066038:
AC 066039:
AC 066040:
AC 066041:
AC 066042:
AC 066043:
AC 066044:
AC 066045:
AC 066046:
AC 066047:
AC 066048:
AC 066049:
AC 066050:
AC 066051:
AC 066052:
AC 066053:
AC 066054:
AC 066055:
AC 066056:
AC 066057:
AC 066058:
AC 066059:
AC 066060:
AC 066061:
AC 066062:
AC 066063:
AC 066064:
AC 066065:
AC 066066:
AC 066067:
AC 066068:
AC 066069:
AC 066070:
AC 066071:
AC 066072:
AC 066073:
AC 066074:
AC 066075:
AC 066076:
AC 066077:
AC 066078:
AC 066079:
AC 066080:
AC 066081:
AC 066082:
AC 066083:
AC 066084:
AC 066085:
AC 066086:
AC 066087:
AC 066088:
AC 066089:
AC 066090:
AC 066091:
AC 066092:
AC 066093:
AC 066094:
AC 066095:
AC 066096:
AC 066097:
AC 066098:
AC 066099:
AC 066100:
AC 066101:
AC 066102:
AC 066103:
AC 066104:
AC 066105:
AC 066106:
AC 066107:
AC 066108:
AC 066109:
AC 066110:
AC 066111:
AC 066112:
AC 066113:
AC 066114:
AC 066115:
AC 066116:
AC 066117:
AC 066118:
AC 066119:
AC 066120:
AC 066121:
AC 066122:
AC 066123:
AC 066124:
AC 066125:
AC 066126:
AC 066127:
AC 066128:
AC 066129:
AC 066130:
AC 066131:
AC 066132:
AC 066133:
AC 066134:
AC 066135:
AC 066136:
AC 066137:
AC 066138:
AC 066139:
AC 066140:
AC 066141:
AC 066142:
AC 066143:
AC 066144:
AC 066145:
AC 066146:
AC 066147:
AC 066148:
AC 066149:
AC 066150:
AC 066151:
AC 066152:
AC 066153:
AC 066154:
AC 066155:
AC 066156:
AC 066157:
AC 066158:
AC 066159:
AC 066160:
AC 066161:
AC 066162:
AC 066163:
AC 066164:
AC 066165:
AC 066166:
AC 066167:
AC 066168:
AC 066169:
AC 066170:
AC 066171:
AC 066172:
AC 066173:
AC 066174:
AC 066175:
AC 066176:
AC 066177:
AC 066178:
AC 066179:
AC 066180:
AC 066181:
AC 066182:
AC 066183:
AC 066184:
AC 066185:
AC 066186:
AC 066187:
AC 066188:
AC 066189:
AC 066190:
AC 066191:
AC 066192:
AC 066193:
AC 066194:
AC 066195:
AC 066196:
AC 066197:
AC 066198:
AC 066199:
AC 066200:
AC 066201:
AC 066202:
AC 066203:
AC 066204:
AC 066205:
AC 066206:
AC 066207:
AC 066208:
AC
```

[illegible]

Mon Dec 21 09:28:30 1998

US-08-951-733-20.rpt

Page 43

```

RP SEQUENCE FROM N.A.
RC STRAIN=2,4,11.
RD MEDIAN: 96082723.
RE HYPOTHESES: 1. H. PARVUS S.;
RF J. BACTERIOL. 177:6422-6431(1995).
RL DBL: U35443; G1072069; -.
RR NON_FER 11 11
R7 13 11
RS SEQUENCE 119 MA: 11864 MW: 08A79MA1 CRM32;

Query Match 0.6%; Score 7; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1,4+e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 45 AARGGA 51
OY 152 ANRGGA 158

RESULT 26 PRELIMINARY: PRT: 117 MA.
ID SOURCE:
AD 030000.
AT 030000.
DT 01-NOV-1996 (TRIMBLER, 01, CREATED)
D7 01-NOV-1996 (TRIMBLER, 01, LAST SEQUENCE UPDATE)
D8 01-NOV-1996 (TRIMBLER, 01, LAST ANNOTATION UPDATE)
DR MICROCATERION LEPARE.
OS PROKARYOTA: FRINGITES: ACTINOMYCETALES: MYCOBACTERIAE.
OC
RN 1) SEQUENCE FROM N.A.
RA SCITE D.B.
RL SUBMITTED (APR-1995) TO EMBL/GENBANK/DDBI DATA BANKS.
RV 12)
RW SEQUENCE FROM N.A.
RX POSITION
RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DDBI DATA BANKS.
RR EMBL: U15183; G699267; -.
RS SEQUENCE 117 MA: 12770 MW: BDAC4C68 CRM32;

Query Match 0.6%; Score 7; DB 2; Length 117;
Best Local Similarity 100.0%; Pred. No. 1,4+e+01;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 65 AARGGS 71
OY 243 ANRGGS 249

```

```

F7 NON_TERR 109 MA: 11728 MW: DDD9548 CRC32:
SQ SEQUENCE 109 MA: 11728 MW: DDD9548 CRC32:
Query Match 0.6% Score 7: DB 3: Length 109:
Best Local Similarity 100.0% Pred. No. 1,47e+01:
Matches 7: Conservative 0: Mismatches 0: Indels 0: Caps 0:
DB 63 CASVAL 69
|||||||
QY 639 CASVAL 705

RESULT 24 PRELIMINARY: PRT: 111 MA.
ID 004233
AC 01-NOV-1986 (TRIMBLER, 01, CREATED)
D 01-NOV-1986 (TRIMBLER, 02, LAST SEQUENCE UPDATE)
D 01-NOV-1986 (TRIMBLER, 02, LAST ANNOTATION UPDATE)
D 01-FEB-1997 (TRIMBLER, 02, LAST ANNOTATION UPDATE)
D8 BIOCHEMICAL 12.3 ED PROTEIN.
CN ONE 1
CN ONE 2 POLYMERASE (TERT)
CC EVOLUTION: PUNG, ASCOCCOPIA, HELIOSCONCITES.
(1)
RC SEQUENCE FROM N.A.
RP STRAIN:DL-1, I. SERATIA, S.A., THORNTON, L.P., ERTHOFF V.M.:
R1 R1
R2 R2
R3 R3
R4 R4
R5 R5
R6 R6
R7 R7
R8 R8
R9 R9
R10 R10
R11 R11
R12 R12
R13 R13
R14 R14
R15 R15
R16 R16
R17 R17
R18 R18
R19 R19
R20 R20
R21 R21
R22 R22
R23 R23
R24 R24
R25 R25
R26 R26
R27 R27
R28 R28
R29 R29
R30 R30
R31 R31
R32 R32
R33 R33
R34 R34
R35 R35
R36 R36
R37 R37
R38 R38
R39 R39
R40 R40
R41 R41
R42 R42
R43 R43
R44 R44
R45 R45
R46 R46
R47 R47
R48 R48
R49 R49
R50 R50
R51 R51
R52 R52
R53 R53
R54 R54
R55 R55
R56 R56
R57 R57
R58 R58
R59 R59
R60 R60
R61 R61
R62 R62
R63 R63
R64 R64
R65 R65
R66 R66
R67 R67
R68 R68
R69 R69
R70 R70
R71 R71
R72 R72
R73 R73
R74 R74
R75 R75
R76 R76
R77 R77
R78 R78
R79 R79
R80 R80
R81 R81
R82 R82
R83 R83
R84 R84
R85 R85
R86 R86
R87 R87
R88 R88
R89 R89
R90 R90
R91 R91
R92 R92
R93 R93
R94 R94
R95 R95
R96 R96
R97 R97
R98 R98
R99 R99
R100 R100
R101 R101
R102 R102
R103 R103
R104 R104
R105 R105
R106 R106
R107 R107
R108 R108
R109 R109
R110 R110
R111 R111
R112 R112
R113 R113
R114 R114
R115 R115
R116 R116
R117 R117
R118 R118
R119 R119
R120 R120
R121 R121
R122 R122
R123 R123
R124 R124
R125 R125
R126 R126
R127 R127
R128 R128
R129 R129
R130 R130
R131 R131
R132 R132
R133 R133
R134 R134
R135 R135
R136 R136
R137 R137
R138 R138
R139 R139
R140 R140
R141 R141
R142 R142
R143 R143
R144 R144
R145 R145
R146 R146
R147 R147
R148 R148
R149 R149
R150 R150
R151 R151
R152 R152
R153 R153
R154 R154
R155 R155
R156 R156
R157 R157
R158 R158
R159 R159
R160 R160
R161 R161
R162 R162
R163 R163
R164 R164
R165 R165
R166 R166
R167 R167
R168 R168
R169 R169
R170 R170
R171 R171
R172 R172
R173 R173
R174 R174
R175 R175
R176 R176
R177 R177
R178 R178
R179 R179
R180 R180
R181 R181
R182 R182
R183 R183
R184 R184
R185 R185
R186 R186
R187 R187
R188 R188
R189 R189
R190 R190
R191 R191
R192 R192
R193 R193
R194 R194
R195 R195
R196 R196
R197 R197
R198 R198
R199 R199
R200 R200
R201 R201
R202 R202
R203 R203
R204 R204
R205 R205
R206 R206
R207 R207
R208 R208
R209 R209
R210 R210
R211 R211
R212 R212
R213 R213
R214 R214
R215 R215
R216 R216
R217 R217
R218 R218
R219 R219
R220 R220
R221 R221
R222 R222
R223 R223
R224 R224
R225 R225
R226 R226
R227 R227
R228 R228
R229 R229
R230 R230
R231 R231
R232 R232
R233 R233
R234 R234
R235 R235
R236 R236
R237 R237
R238 R238
R239 R239
R240 R240
R241 R241
R242 R242
R243 R243
R244 R244
R245 R245
R246 R246
R247 R247
R248 R248
R249 R249
R250 R250
R251 R251
R252 R252
R253 R253
R254 R254
R255 R255
R256 R256
R257 R257
R258 R258
R259 R259
R260 R260
R261 R261
R262 R262
R263 R263
R264 R264
R265 R265
R266 R266
R267 R267
R268 R268
R269 R269
R270 R270
R271 R271
R272 R272
R273 R273
R274 R274
R275 R275
R276 R276
R277 R277
R278 R278
R279 R279
R280 R280
R281 R281
R282 R282
R283 R283
R284 R284
R285 R285
R286 R286
R287 R287
R288 R288
R289 R289
R290 R290
R291 R291
R292 R292
R293 R293
R294 R294
R295 R295
R296 R296
R297 R297
R298 R298
R299 R299
R300 R300
R301 R301
R302 R302
R303 R303
R304 R304
R305 R305
R306 R306
R307 R307
R308 R308
R309 R309
R310 R310
R311 R311
R312 R312
R313 R313
R314 R314
R315 R315
R316 R316
R317 R317
R318 R318
R319 R319
R320 R320
R321 R321
R322 R322
R323 R323
R324 R324
R325 R325
R326 R326
R327 R327
R328 R328
R329 R329
R330 R330
R331 R331
R332 R332
R333 R333
R334 R334
R335 R335
R336 R336
R337 R337
R338 R338
R339 R339
R340 R340
R341 R341
R342 R342
R343 R343
R344 R344
R345 R345
R346 R346
R347 R347
R348 R348
R349 R349
R350 R350
R351 R351
R352 R352
R353 R353
R354 R354
R355 R355
R356 R356
R357 R357
R358 R358
R359 R359
R360 R360
R361 R361
R362 R362
R363 R363
R364 R364
R365 R365
R366 R366
R367 R367
R368 R368
R369 R369
R370 R370
R371 R371
R372 R372
R373 R373
R374 R374
R375 R375
R376 R376
R377 R377
R378 R378
R379 R379
R380 R380
R381 R381
R382 R382
R383 R383
R384 R384
R385 R385
R386 R386
R387 R387
R388 R388
R389 R389
R390 R390
R391 R391
R392 R392
R393 R393
R394 R394
R395 R395
R396 R396
R397 R397
R398 R398
R399 R399
R400 R400
R401 R401
R402 R402
R403 R403
R404 R404
R405 R405
R406 R406
R407 R407
R408 R408
R409 R409
R410 R410
R411 R411
R412 R412
R413 R413
R414 R414
R415 R415
R416 R416
R417 R417
R418 R418
R419 R419
R420 R420
R421 R421
R422 R422
R423 R423
R424 R424
R425 R425
R426 R426
R427 R427
R428 R428
R429 R429
R430 R430
R431 R431
R432 R432
R433 R433
R434 R434
R435 R43
```

Mon Dec 21 09:28:30 1998

US-08-951-733-20.rpt

Page 44

```

D7 01-AGC-1898 (REBLENED.. 07, CRAFTED)
D7 01-AGC-1898 (REBLENED.. 07, LAST SEQUENCE UPDATE)
D7 01-AGC-1898 (REBLENED.. 07, LAST ANNOTATION UPDATE)
D7 GAG PROTEIN (FLOMAGENT)..
G3
G3 GAG
OS HUMAN IMMUNODEFICIENCY VIRUS TYPE 2 (HIV-2).
OS VARIANTE; SS-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRIDAE;
OC LENTIVIRINE.
NC [1]
RP SEQUENCE FROM N.A.
RA STRAIN=4.. ANTIOSSI K., MILKINS A., DIAS F., WEITLIE B., BREUER J.;
RA XIANE S. NON. RETROVIRUSES 11:501-506(1997).
RH [12]
RP SEQUENCE FROM N.A.
RC STRAIN=4.. XIANE S., ANTIOSSI K., MAET P., JENSEN H., DIAS F.,
RA WEITLIE B., BREUER J.,
RA SUBMITTED (MAR-1998) TO EMBL/GENBANK/DDBB DATA BANKS.
DR DBEM, LA008512, E1297564.
LC NON-TEAR 125 125
SC SEQUENCE 135 AA; 13505 MW; FC040DC CMCR32;
Query Match
Score: 10.68; Score 2: DB 14; Length 135;
Identical Similarity 100%; Identical 0; Mismatches 0; Indels 0; Gaps 0
Matches 7; Conservative 0
Db 7 PERSEP RP 13
QY 330 PERSEP 336
RESULT 38
AC Q55981 PRELIMINARY: PRP: 128 AA.
AC Q55981.
D7 01-NOV-1998 (REBLENED.. 01, CRAFTED)
D7 01-NOV-1998 (REBLENED.. 01, LAST SEQUENCE UPDATE)
D7 01-NOV-1998 (REBLENED.. 01, LAST ANNOTATION UPDATE)
D2 RPHOCTHEIN 14.4 KD PROTEIN
G3 SILO651.
OS STREPTOCOCCUS SP.
OS STREPTOCOCCUS SP.
OS STREPTOCOCCALIS.
OS CHROMOCOCCALIS.
NC [1]
RP SEQUENCE FROM N.A.
RA STRAIN=PCC6803.
RA NON-TEAR 125 125
RA KAMETO T., TANAKA A., SATO S., KOTANI H., SAOTIA T., MIYAHARA N.,
LA SUJIRA N.H., ZABAYT S.;
LA DNA RES. 2:153-166(1995).
```

DR EMBL D64005: G1006407: -
 NC NUCLEOTIDE SEQUENCE: 188 AA: 14443 MW: 20040558 CRC32:
 SQ SEQUENCE 149 AA: 14443 MW: 20040558 CRC32:
 Query Match 0.64: Score 7; DB 2; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 46 AVVIDOS 52
 |||||
 QY 811 AVVIDOS 817

RESULT 29
 ID 080839 PRELIMINARY: PRT: 135 AA.
 AC 080839: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 135 AA: 14124 MW: 18954028 CRC32:
 RX MEDLINE: 92260186.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 135 AA: 14124 MW: 18954028 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 72 SRSIPLP 78
 |||||
 QY 251 SRSIPLP 257

RESULT 30
 ID 080833 PRELIMINARY: PRT: 138 AA.
 AC 080833: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 138 AA: 14124 MW: 18954028 CRC32:
 RX MEDLINE: 92260186.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 138 AA: 14124 MW: 18954028 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 135;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 72 SRSIPLP 78
 |||||
 QY 251 SRSIPLP 257

ID 082408 PRELIMINARY: PRT: 149 AA.
 AC 082408: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 149 AA: 15702 MW: 4830545 CRC32:
 RX MEDLINE: 96190544.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 149 AA: 15702 MW: 4830545 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257

RESULT 33
 ID 082405 PRELIMINARY: PRT: 149 AA.
 AC 082405: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 149 AA: 15702 MW: 4830545 CRC32:
 RX MEDLINE: 96190544.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 149 AA: 15702 MW: 4830545 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257

OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 149 AA: 15702 MW: 4830545 CRC32:
 RX MEDLINE: 96190544.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 149 AA: 15702 MW: 4830545 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257

RESULT 31
 ID 046381 PRELIMINARY: PRT: 146 AA.
 AC 046381: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 146 AA: 14612 MW: 53345260 CRC32:
 RX MEDLINE: 92260186.
 RA EMBL D63524: G940401: -
 DR EMBL D63524: G940401: -
 FT NON_TER 1
 SQ SEQUENCE 146 AA: 14612 MW: 53345260 CRC32:
 Query Match 0.64: Score 7; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 58 LABELING 64
 |||||
 QY 372 LABELING 378

RESULT 32
 ID 046381 PRELIMINARY: PRT: 146 AA.
 AC 046381: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 146 AA: 14612 MW: 53345260 CRC32:
 RX MEDLINE: 92260186.
 RA EMBL D63524: G940401: -
 DR EMBL D63524: G940401: -
 FT NON_TER 1
 SQ SEQUENCE 146 AA: 14612 MW: 53345260 CRC32:
 Query Match 0.64: Score 7; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 58 LABELING 64
 |||||
 QY 372 LABELING 378

RA EMBL D63524: G940401: -
 DR EMBL D63524: G940401: -
 FT NON_TER 1
 SQ SEQUENCE 146 AA: 14612 MW: 53345260 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257

RESULT 34
 ID 082387 PRELIMINARY: PRT: 149 AA.
 AC 082387: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 149 AA: 15702 MW: 4830545 CRC32:
 RX MEDLINE: 96190544.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 149 AA: 15702 MW: 4830545 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257

RESULT 35
 ID 082405 PRELIMINARY: PRT: 149 AA.
 AC 082405: -
 DT 01-NOV-1996 (TRIMBLAETL. 01, CREATED)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DT 01-NOV-1996 (TRIMBLAETL. 01, LAST SEQUENCE UPDATE)
 DE T-CELL LYMPHOTOXIC VIRUS TYPE II (CLONE NO-6) TAX-2 (FRAGMENT).
 OS HUMAN T-CELL LYMPHOTOXIC VIRUS TYPE II.
 OC RETROVIRIDAE: ONCOTRINAE: TYPE C ONCOVIRUS GROUP.
 NC NUCLEOTIDE SEQUENCE: 149 AA: 15702 MW: 4830545 CRC32:
 RX MEDLINE: 96190544.
 RA EMBL D63885: G139556: -
 DR EMBL D63885: G139556: -
 FT NON_TER 1
 SQ SEQUENCE 149 AA: 15702 MW: 4830545 CRC32:
 Query Match 0.64: Score 7; DB 14; Length 149;
 Best Local Similarity 100.0%; Pred. No. 1.47e+01;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Db 86 SRSIPLP 92
 |||||
 QY 251 SRSIPLP 257


```

DR      EMBL: J13051; E1166501;
SQ      SEQUENCE 170 AA: 18402 MW: 38635442 CRC32:
Query Match      0 6%: Score 7: DB 14%: Length 170:
Beat Local Similarity 100.0%: Pred No. 1,476+0!:
Matches          0: Conservative          0: Mismatches 0: Indels 0: Gaps 0:
Db              107 SRS1PLP 113
              251 SRS1PLP 257

RESULT 42
ID      085601      PRELIMINARY:      PRT:      170 AA.
DC      085601:1986 (TREMURER, 01 CREATED)
DD      D.D.
DT      01-NOV-1996 (TREMURER, 01, LAST SEQUENCE UPDATE)
DT      01-NOV-1996 (TREMURER, 01, LAST SEQUENCE UPDATE)
DD      REX 36 KD PROTEIN
DE      36 KD PROTEIN
DE      VIRALDE: SS-RNA NONENVELOPED VIRUSES, POSITIVE STRAND RNA VIRUSES:
OC      RETROVIRIDAE: ONCOVIRINAE: TYPE C ONCIVIRUS GROUP:
CC      MAMMALIAN TYPE C ONCIVIRUSES.
NC
DC      130 SEQUENCE FROM N.A.
DD      D.D.
RX      MEDLINE: 84250189
BA      SODOSQ J., PINKAS R., PERKINS D., BRIGGS D., LEE T.H., ESSEX M.,
BA      COLGAN J., WONG-STAL P., CALVO R.C., MASSUTTE M.A.,
BA      SODOSQ J. 1994:321-441(1994).
RX      MEDLINE: 85216448
BA      130 SEQUENCE FROM N.A.
DD      D.D.
RX      CHEN I.C., KIVA M., SHIMIZU Y., SHIMIZU N., OGUCHIYAMA T., GOLDBE D.N.,
RX      PROC. NATL. ACD. SCI. U.S.A. 82:3101-3105(1985).
NR      (1)
NR      130 SEQUENCE FROM N.A.
DD      D.D.
NR      (1) WEISS R.A., VARGAS R., COFFIN J.:
NR      (1) WEISS R.A., VARGAS R., COFFIN J.: (EMSA.):
NR      RNA TUMOR VIRUSES, SECOND EDITION, 2, VOL. 2:1070-1085;
NR      COLD SPRING HARBOR LABORATORY, COLD SPRING HARBOR (1985).
NR      130 SEQUENCE FROM N.A.
DD      D.D.
BA      ROSENBLATT J.:
BA      SUBMITTED NOV-1986 TO EMBL/GENBANK/DBJ DATA BANKS.
NR      EMBL: M10060; G139360;
NR      130 SEQUENCE 170 AA: 14663 MW: 00011824 CRC32:
Query Match      0 6%: Score 7: DB 14%: Length 170:
Beat Local Similarity 100.0%: Pred No. 1,476+0!:
Matches          0: Conservative          0: Mismatches 0: Indels 0: Gaps 0:
Db              107 SRS1PLP 113
              251 SRS1PLP 257

```

```

US-08-951-733-20.rpt
US Dec 21 09:28:30 1998

DB      107 SHEDUP 113
  QY      251 SHEDUP 257

      RESUTR      43
      ID      068704      PRELIMINARY;      PRT:      170 AA.
      AC      008838;
      DT      01-NOV-1996 (TRIMDATEL, 01, CREATED)
      DT      01-NOV-1996 (TRIMDATEL, 01, LAST SEQUENCE UPDATE)
      DT      01-NOV-1998 (TRIMDATEL, 02, LAST ANNOTATION UPDATE)
      DB      REX PROTEIN
      CD      REX
      CG      HUMAN T-CELL LEUKEMIA VIRUS TYPE II (ISOLATES G12 AND MR4) (HTLV-II).
      CC      VIRALDE; S-RNA ENVELOPED VIRUSES; POSITIVE-STRAND; RETROVIRINE.
      RN      113
      RP      SEQUENCE FROM N.A.
      RA      MEDLINE; 93323207.
      RA      FORD D., SMITHER M. M., BALLOCK K. G., KAPLAN J. E., LAL R. B.,
      RE      J. VIROL. 67:4659-4664(1993).
      RN      [2]
      RP      SEQUENCE FROM N.A.
      RA      MEDLINE; D3366434.
      RA      SMITH R. F., SHANSON P., ABALICIO J. J., CHIN K. K., LAY J. P.,
      RA      NOTTER M., MANN T., LENCZE G., ANNETTI A., MARINCCI G., CHEN I. S. T.,
      RA      ROSENBLATT J. D.;
      RE      VIROLOGY 196:57-69(1993).
      DB      REX; L0735; G04041
      DB      REX; L0735; G04041
      SQ      SEQUENCE      110 AA; 18372 MW; 66246EE6 CRC32;

Query Match      0.64; Score 7; DB 14; Length 170;
      Best-Local Similarity 100.0%; Pctd. No. 14/e=0.1;
      Matches      0; ConservedAtc      0; Identical      0; Incls      0; Gaps      0
      Db      107 SHEDUP 113
      QY      251 SHEDUP 257

      PRELIMINARY;      PRT:      182 AA.
      ID      068704
      AC      018902;
      DT      01-NOV-1998 (TRIMDATEL, 07, CREATED)
      DT      01-NOV-1998 (TRIMDATEL, 07, LAST SEQUENCE UPDATE)
      DT      01-NOV-1998 (TRIMDATEL, 07, LAST ANNOTATION UPDATE)
      DB      YOPF (YOPF) HOMOLOG (YOPF).
      CD      YOPF/YOPF.

```

Mon Dec 21 09:28:30 1998

US-08-951-733-20-Report

Page 55

US-08-951-733-20.rpt

Page 56

[illegible]

```

Db      62 GS18LT 68
Oy      856 GS18LT 862

RESULT 16
AC 062143 PRELIMINARY; PRT: 164 AA.
D1 01-NOV-1996 (THERMSELT. 01, CREATD)
D7 01-NOV-1996 (THERMSELT. 01, LAST SEQUENCE UPDATE)
D8 08/2D CASE8 (THERMSELT. 05, LAST ANNOTATION UPDATE)
OS MOS MOSCUDUS (MOSCS).
OC ERYTHROCA. METACOA. CORDATA. VERTEBRATA. TETRAPODA. MAMMALIA.
CC ERYTHRIA. ROSEBATA.
EP SEQUENCE FROM N.A.
RA NAKAYAMA K.I., TOKITO S., NAKADUCHI H., GACHELIN G.;
RL IMMOBILIZATION: 0-0-0-0.
DR PNA: P001139. HNC I.
SQ SEQUENCE 184 AA: 21191 MW, 8545770D CRC32;

Query Match 0.64 Score 7, DB 11, Length 164;
Query Local 7; Conservative 0; Matches 0; Indels 0; Gaps 0
Db      97 GS18AT 103
Oy      1116 GS18AT 1122

RESULT 17
AC P71395 PRELIMINARY; PRT: 192 AA.
D7 01-FEB-1997 (THERMSELT. 02, CREATD)
D7 01-FEB-1997 (THERMSELT. 02, LAST SEQUENCE UPDATE)
D7 01-FEB-1998 (THERMSELT. 07, LAST ANNOTATION UPDATE)
D8 8/1883.
OS HEMIPHILUS INTERSEAE.
OC PROKARIOTA. GINCHILICUTES. SCOTOBACTERIA. FACULTATIVELY ANAEROBIC ROOS;
CC HEMIPHILUS.
EP SEQUENCE FROM N.A.
RA MEDLINE, 93550630.
RL KIRKNESS R.B., BUTT C.J., TOMB J.F., DOUGHERTY B.A., HENRICK J.M.,
RA KIRKNESS R.B., BUTT C.J., TOMB J.F., DOUGHERTY B.A., HENRICK J.M.,
RA MENNERIKY K.B., STITTON G., FITZROGH W., FIELDS C.A., GOCVART J.D.,
RA SCOTT J.D., SHIRLEY R., LIO L.I., GLOVER N., KETLER J.M., WEIDMAN J.F.,

```


THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)